

上海译文出版社

Roger S. Frantz

X EFFICIENCY: THEORY, EVIDENCE AND APPLICATIONS

Kluwer Academic Publishers 1988

根据美国克卢渥学术出版公司 1988 年版译出

X效率：理论、论据和应用

〔美〕罗杰·S·弗朗茨 著

费方域等译

上海译文出版社出版、发行

上海延安中路 955 弄 14 号

全国新华书店经销

上海译文印刷厂印刷

开本 850×1168 1/32 印张 6.75 插页 2 字数 [illegible]

1993 年11月第1版 1993 年11月第1次印刷

印数：0,001—3,000 册

ISBN7-5327-1315-8/F·033

定价：7.70元

(沪)新登字111号

新古典经济学和X效率理论
——译序——

X(低)效率理论，从哈维·莱宾斯坦1966年最早提出迄今，一直在积极尝试批判性地重构经济学中长期占统治地位的新古典学说，尤其是它的厂商理论。

新古典学派把经济学看作研究稀缺资源通过市场在各种可供选择的用途间进行有效配置的学问。因此，它主要关注市场(价格)的配置效率——帕累托效率——及其实现条件，而完全疏漏了企业（个人和集体）内部效率——X效率——及其实现条件。

新古典理论假设，厂商是根据生产函数和成本函数进行生产的，换言之，就是厂商总是在既定投入和技术水平下实现产量极大化和单位成本极小化的。投入产出(或成本产出)关系，只是一种与每个人的决策行为无关的纯技术关系。因此，这个理论从一开始，就通过假设把出现X低效率的可能性排除掉了。

但是，X(低)效率的存在却是一个客观事实。它实质上是一种组织(organizational)或动机(motivational)(低)效率。因此，这个X代表造成非配置(低)效率的一切因素。在厂商要素投入量给定的情况下，如果要素投入变化可以使得某些产品的产出增加而没有其他产品的产出减少，就说明该厂商存在X低效率，反之，则说明它有X效率。

当初，莱宾斯坦就是因为看到以下一些与新古典理论描述的企业营运方式完全相违的事实：(1)企业不是内部有效率的，即只要将工厂的内部组织作简单的变动，厂商就能增加它们的产量；(2)企业不是利润极大化的，即厂商并不按边际分析原理经营；(3)存在着劳动和资本以外的某种东西在工业化国家的增长率中发挥重要作用，才认为有必要提出一个新的概念，用来归纳这一类普遍存在的非配置型低效率现象。因为这种形式低效率的性质当时尚不明了，所以莱宾斯坦把它称作X低效率。

传统新古典理论与实际存在的X(低)效率现象间的不一致性，标志着传统新古典理论陷入了危机。对于有为的经济学家来说，这是一次发展经济理论的新契机，他们当然不会轻易错过它。

一部分经济学家，如施蒂格勒、波斯纳、塔洛

克等人，在继续坚持新古典基本原理和方法的前提下，通过修正某些假设(如将利润最大化改为效用极大化)或引入新的分析环节(如产权、交易费用、寻租等等)的办法，努力使原来的分析框架也能处理企业内部存在的机会未被利用的这类低效率现象。在这部分人看来，理论假设只是一种使问题得到简化的抽象，它们既不必被直接观察到，也不必与可观察到的现象直接相符一致，只要能在假设和现象间建立起合乎逻辑的联系，假设及由其推出的结论就算是有效的了。

相反，另一部分经济学家，如莱宾斯坦和他的追随者们则认为，假设也应该是可以被直接观察到的。因此，基本假设不完全现实的新古典经济学，特别是它的厂商理论，必须重新构造，以使经济理论不仅能说明市场配置效率，而且也能说明企业的X效率。

因此，从前提、方法、构架直至结论，后者都提出了自己独特的看法，从而形成了区别于传统新古典经济学的X效率理论。

在新古典的传统观点看来，第一，代表性企业是生产者的恰当研究单位，代表性家庭是消费者的恰当研究单位。第二，无论是生产者还是消费者，都

是理性的经济人，他们的行为，都是实现极大化的行为，比如，企业的行为是追求利润（或其他目标）极大化，家庭的行为是追求效用极大化。第三，在企业内部，雇员的行为就是为了使企业主的目标（比如说利润）实现极大化，换言之，雇员没有自己的与企业的目标和利益不相一致甚至冲突的目标和利益。第四，经济行为人肯定会对环境的变化作出充分的反应。第五，劳动合同是完整的，这就是说，不仅对于雇员的报酬是确定的，而且雇员的劳动时间和努力程度也是确定的。第六，雇员（经济和其他人）的努力程度是一个既定的常量。

针锋相对地，X（低）效率理论提出了一整套他们认为更实际、适用范围更广泛的假设和命题。

第一，生产活动不只是可以借助现代数学和物理方法精确描述的一种技术决定关系，在一定程度上，它决定了个人的心理和生理活动。由于只有个体的人才有思想和行动，所以，作为经济学研究对象的恰当单位就不应该是笼统的家庭和企业，而应该是构成它们的最小行为单位——个人。

第二，任何个人都具有双重性。即个人的行为，一方面，具有确定和坚持标准，通过计算和注意细节，努力追求极大化的倾向；另一方面，又具有不

关心标准，不注意计算和细节，不努力追求极大化的倾向。这两种倾向的对立和并存，决定了新古典所谓的完全理性的经济人只能是一种极端的和个别的情况：是某些人在某些时候某些条件下采取的行为的特征，而不是所有的人在所有时候所有条件下采取的行为的特征。通常的情况应该是，个人只具有有选择的理性，他们只把一部分精力放在作出信息充分的决策所必需的细节和计算上，而让更多的决策采取依赖于习俗、惯例、道德规范、标准程序和模仿的形式作出(这些形式明显具有非极大化的特征)。这就是说，个人的行为应该既包含理性因素，又包含非理性因素。因此，应该用极大化非极大化假设来替代新古典经济学的极大化假设：不排除极大化行为，但是，也必须允许、承认、考虑和研究非极大化行为。对于理解X效率理论和新古典理论的不同来说，这个区别至关重要。

为了能统一处理极大化和非极大化行为，X效率论者利用并改写了耶基斯—多森心理法则。这个法则在1908年由实验心理学家罗伯特·耶基斯和约翰·多森提出的时候，揭示的是刺激强度与学习的关系。莱宾斯坦在他1983年和1985年的两篇著述中，将其借用来表现压力和工作绩效之间的二次关

系：承受相对较高和较低压力的个人，是不会努力对决策作仔细计算，不会尽可能做好工作的，只有在适度的压力下，他们才能采取极大化行动，工作才可能有最好的绩效。值得注意的是，这里明确引入了压力这个概念，并表述了它与X（低）效率之间的关系。这个压力，既可以由个人自身内部机制创造出来，例如，有的人比别人更倾向于按责任、义务和标准行事；也可以由外部环境通过企业内部人际关系（上司，同事）和市场竞争（或缺乏竞争）的效应形成，并且，可以表现为不同的强度。这个概念，是新古典经济理论没有采用的。

第三，应该把个人的努力（体力和脑力的运用），看作是个人对他（或她）自己精神和（或）由外部环境确定的动机作出反应的结果，努力的构成要素，至少有活动(A)、进度(P)、质量(Q)和时间模式(T)四种，即所谓的APQT束。任何个人，对APQT束都有一定程度的自由选择权。这种选择，不仅根据成本和收益（新古典就强调这一点），而且，还要根据动机和认识系统（X效率理论将经济行为分析引向它的心理学基础）。因此，由于(1)个人并不一定有足够的动机按照极大化行为模式来思考问题；(2)即使个人确实在按照偏好进行选择，习惯也会对这些

偏好和选择作出调节；(3) 个人即使有偏好，也不一定有能力或意愿对所处环境的各种变化都作出区分，这也就是说，他(她)们并不总是知道什么时候可以对什么样的两种不同的状况作出选择，所以，个人的努力程度不应该是一个机械地决定的常量，而应该是一个任意决定的变量。

第四，个人的行为具有惰性特征，或者说，个人通常都在他(她)的“惰性区域”工作。这种所谓的“惰性区域”，实际上是一个努力区域，在这区域的下限以下，随着努力水平的提高，个人的满足程度会以递增的速率增加，在这区域的上限以上，随着努力水平的提高，个人的满足水平会以递减的速率下降，因此，只有这个区域中的努力水平，才是令人满意的“均衡”努力水平。在这个区域中，各个努力水平、各个满足水平彼此间很难加以区分。个人一旦进入工作的惰性区域，就会抵制改变其努力水平的各种影响，因此，自变量变动不导致应变量相应变动，便成为这个区域最重要的特征。这就再一次证明，个人只有有选择的理性：他们并不总是竭尽所能来计算和分析各种不同努力和满足水平的成本和利益的。

第五，生产过程不是一种机械过程，企业也不

是一部将投入变为产出的高效转换器。X效率的实现要依赖企业全体成员的努力。在企业中，上司、同事和传统都会对个人的努力水平发生影响。通常，来自上司的压力(如监督和奖惩)会使个人更加勤快地工作，即变得更加理性。经理们只要创造一个能使雇员感到他们是在为他们自己工作的环境，并使雇员们感到他们在工作中是与他人平等的合伙者，就能提高雇员们的努力水平。而来自同事的压力，却可能产生三种不同的效应：(1) 大家都在尽职尽力，你也要尽职尽力；(2) 如果你想偷懒，你可以偷懒，但却不可以太卖力，以免显出别人不卖力；(3) 尽你自己的职责，但不可以太努力，免得使别人显得不努力。至于传统的影响，可以非常典型地在这样一则对答中反映出来："你为什么这样工作？""不知道，我们一直就是这样做的。"在一个拥有众多成员的企业中，长期形成的集体工作规范，会使得集体表现出的努力水平接近于每个集体成员的"平均"生产率水平。

第六，劳动合同是不完善的。在雇佣合同中，企业直接购买的是劳动时间，而不是作为人力投入最重要因素的劳动努力，但生产使用的却是劳动努力。企业要事先为雇佣人员确定他们的APQT束，涉

及的搜寻成本、监督成本是十分高昂的，甚至会导致后者怠工和旷工。所以，这种合同安排总是不完全的。按照这种合同，雇员对于提供多少某种技能水平上的努力，具有相当的自由处置权。因此，雇员的实际努力程度，还必须取决于企业的激励机制。通过激励机制，压力影响动机，动机影响努力，努力影响成本，从而提高X效率，或者，降低X低效率。劳动投入不同于其他投入，是X效率理论的一条重要立论依据。

第七，企业主（委托人）和雇员（代理人）的利益并不总是一致的。企业内决定劳动程度的所有变量，并不由企业主全部控制，而是如上所述，一部分由雇员控制，另一部分由雇主控制。因此，企业生产率的决定问题，就成了一个典型的决策论问题。莱宾斯坦用囚徒的两难困境对策来描述这个问题：企业和雇员都有三种策略可供选择。其中，企业的极大化策略是使提供给雇员的工作条件和工资极小化，而同时又试图使能从雇员那里得到的东西达到极大。雇员的极大化策略是尽可能小地付出努力，但却要照拿工资，采取这种策略的每个雇员都以为，他只是免费搭车，他个人的偷懒不会使企业遭受损失从而引起注意，所以也不会使他的个人利益（如

收入和岗位)受到影响。可是，如果双方都采取极大化策略的话，就会导致囚徒的两难困境——一个生产率低下，报酬和工作条件可怜的企业。另一方面，企业可以采取完全是为了职工利益（不管职工表现如何)的黄金法则策略，或者，雇员也可以采取为企业最大利益工作(不管企业如何对待他们)的黄金法则策略，但是，除非双方同时采取这种策略，不然，采取这种策略的任何一方都会感到自己由于对方未采取这种策略而陷于"不公平"的境地。最后，如果双方都采取选择同业集体标准的策略，企业按习惯提供工作条件(包括收入)，雇员按习惯提供工作努力，不合习惯的一方或者调整自己的行为，或者另谋他就以维持这种合作，那么，就像所有汽车都沿公路一侧行驶的习惯能解决到底汽车该沿公路左侧还是右侧行驶的选择问题一样，同业集体标准解也能使双方的付出和收入相等，因此，这种合作解优于其他所有的非合作解。

第八，市场结构与X(低)效率有关，垄断企业不仅会造成低配置效率(由生产者剩余和消费者剩余表示的净经济福利损失)，而且会造成缺乏动机、有机会不利用的X低效率。由于垄断企业可以通过控制价格，无需成本极小化就能实现可以接受的利

润水平，所以，如果雇员有任意决定努力水平的自由的话，垄断力量就会使他们与任何降低成本的压力隔绝。只要垄断企业满意它的利润约束，那就没有理由认为成本极小化是它的典型行为。只有完全竞争的市场，才会经过价格机制，对企业产生不降低成本和价格就要从行业中淘汰出去的压力，从而增加企业成员的危机感、责任感和成就感，克服惰性，提高努力水平。

第九，在减少X效率方面，企业家的才能发挥重要作用：或者对低效率企业直接发生影响；或者作为竞争者对企业发生间接影响。在不完全市场上，这种作用尤为明显。所谓不完全市场，一是指安排和利用投入存在障碍，如有些需要的投入正被用于其他方面，有些投入价格不合适，有些投入不是人人都弄得到等等；二是指投入产出市场存在各种漏洞，如知识并不总是可以交易的，把知识从想法变为可用形式需要激励等等，企业家在不完全市场中的作用，就是克服这些障碍，填补这些漏洞。

新古典学派和X效率学派之间的是非曲直，现在正争得难分难解。对于我们来说，重要的不仅在于判别他们的孰是孰非，而且还在于从双方的最新研究成果中吸取对我们有用的东西。联系到当前的

改革实践，我想，我们至少可以因此认识到，第一，既要重视配置效率，又要重视X效率；第二，政企不分，产权虚设，缺乏竞争，预算软约束等等弊端，既影响配置效率，又影响X效率；第三，向市场经济过渡，转换企业运行机制，不仅能提高配置效率，而且能提高X效率。

最后需要说明的是，译者在翻译过程中请马胜同志帮助翻译了本书的第7、8章，王丽琼同志翻译了第9、10章。上海译文出版社的张宇宏同志做了不少工作，特在此一并表示感谢。

费方域

1992.10.8

前言

我对X效率发生兴趣，可以追溯到1978年。当时，我正在华盛顿州立大学写博士论文。论文的内容是关于青年人在从学校到工作过渡时期中所抱态度的作用。米拉德·黑斯蒂(Millard Hastay)教授(本人主持的委员会的一位成员)劝我读一下莱宾斯坦(Leibenstein)的"新"著《超越经济人》(Beyond Economic Man)，这本书对于理性(选择)行为的描述，引起了我的注意。看来，莱宾斯坦对于理性个人的选择行为的描述，与亚伯拉罕·马斯洛那样的心理学家所报道的非常相似，都说这种行为是特殊动机系统的产物。换言之，我得到的印象是，莱宾斯坦称为X低效率的东西，在心理学家那里，是作为"它(经常)就是那个样子"的东西讨论的。因此从一开始，我就一直认为X(低)效率对于理解人的行为是一个有价值的概念。自那时起，我越来越相信，在考虑非市场环境——即企业——中的行为时，这一点尤其正确。

然而，确切地说，本书的研究工作是在1982年开始的。当时，我在哈佛大学当访问学者，莱宾斯坦教授建议，让我探究被用作证明产权作用的某个经验证据如何也能与X效率理论相一致(双方的一致性是相当多的)。由此导致了对把X效率

作为报告结果的一种可能解释的经验研究进行的一次调查。这方面的文献看来很多，但不像引用X效率的理论文献那么多。我把这大量的文献搜集到一起，加以编目，其中的一个方面，我的许多同事和其他我与之说起过此事的人都不知晓。我的印象是，经济学界有许多人还不知道论X效率理论的这部分文献。

给这本书最大帮助的是哈维·莱宾斯坦。他为本书和其他研究项目，慷慨地花费了大量时间，给予了我巨大的鼓励。我了解他是一个不惧怕在正统思想框架外进行思索的人。他敏锐的、富有启迪和教育意义的洞察力，给我留下了深刻的印象。虽然莱宾斯坦和我在某些问题上可能意见不一致，但我相信，在本书一开头就声明"我是一个X主义者"是顺乎情理的。

还有许多人读过此书的一些章节，或者，为这项研究的进展贡献过他们的时间和专业知识。我特别要感谢：迈克尔·克鲁(拉特格斯大学)，他是包括本书在内的一套丛书的编辑；琼·安德森(圣迭戈大学)，我和她一起研究过一些课题；理查德·哈特威克(西伊利诺伊大学)，他是《行为经济学杂志》的编辑，他特别支持我的一些想法。还要感谢唐·勒克劳(西安大略)、沃尔特·普里莫克斯(伊利诺伊大学)和T·Y·申(加利福尼亚大学戴维斯分校)。

我在加州圣迭戈大学的几位同事也曾慷慨地为我奉献过他们的时间和专业知识。哈林德·辛格几乎通读了我的手稿，并提供很有价值的评论和组织材料的思路。迈克尔·诺顿指导过我管制经济学。当然本书在这个领域的不足只由我自己负责。我想表示谢意的其他同事还有乔治·巴比洛、雷·

博迪、弗雷德·盖洛韦、卢·格林、易卜拉欣 波罗伊、弗雷德·西博尔德、罗伯特·塞德曼和伊安尼斯·韦尼耶斯。

我还要感谢我的朋友奥利亚诺·克鲁斯，他使我有可能在墨西哥研究X效率，没有他，在墨西哥的研究工作无法完成。最后，还有朱迪·巴克（他是我第10章所描绘的“发疯”顾客的原型）和马克·克鲁克斯兄弟，他们为本书发挥了极好的美术才能。

加州圣迭戈大学公共经济学中心，为我研究这个课题提供了财政支持，并于1984年出版了我的初稿。本书就是由这个初稿扩写而成的。我感谢该中心主任乔治·巴比洛为这个研究项目的进展提供的巨大支持。本书第9章“X效率：它的批评者和答复”最先发表在《经济学和商业季度评论》上，我感谢该杂志编者保罗·尤塞尔丁允许我随意地借用这篇文章。

罗杰·S·弗朗茨

序言

大多数基础经济学的教师都会碰到不相信他们所教的一些理论的学生。学生会争辩说:"在真正的企业中不是这样做的",或者"我爸爸的公司里不是这样做的"。为了对付这种挑战,并向他们的学生说明标准理论的合理性,经济学教师常常给予这样一些方便的回答:"要理解经济运行这样复杂的现象,需要一种简化的模型","所有科学都进行简化","如果我们对它理解更加彻底的话,看来似乎不正常的行为就不会觉得异常了","我们不能一下子从全部复杂现象着手"。

在某种意义上,有没有这种可能,即这些辩护理由是错的,而学生的反对意见却是对的呢?即使简化是必须的,我们经济学家难道就能够争辩说,我们已经真正找到了一整套建立我们所希望建立的那种有效并且有用的经济学的简化假设吗?这是我关注了几十年的问题。在这篇序文中,我要尽可能对这个普遍的问题以及几个相关的问题进行多侧面的深入考察。

我对X效率的研究基本上是一件偶然事情——没有充分使用我的研究助手——的结果;他们愿意找出大部分由国际劳工组织(ILO)和联合国人员访问欠发达国家的企业时写出

的专门报告的细节，这些报告收藏在加利福尼亚大学图书馆里，他们工作揭示出的许多鲜明的实际例子，说明有许多企业的经营不是依照最优的方式，而是依照与标准微观理论相矛盾的方式。鉴于这些资料，我不得不重新考虑先前所持的立场。1966 年最早论X效率的那篇文章就是这些考虑的结果。我并没有想到该文会被大量引用，促进别人的研究工作以及引起一些论战。

对X效率理论的普遍反应引出了关于现行微观经济学的性质的问题。微观经济学是一门累积性的科学吗？它完全是以渐进的方式建立在过去成就的基础上的吗？目前它的状况就是我们的知识一点一点增长的结果吗?它是逐步发展的吗?总而言之，微观经济学有过两次重大"革命"：1776年的斯密学派革命，和19世纪70年代的边际主义革命（其中包括瓦尔拉的著作)。第一场革命出现在二百多年以前,第二场革命则发生在一个多世纪以前。19世纪70年代以来，研究工作似乎只有微小变化的进展，大部分也只是使理论的数学解释变得更加精细优美。

经济学家们的普遍看法是，这种精益求精的过程对理解尤其是精确的理解有好处。但是凡事有得必有失。部分的精细化过程会把理论中似乎是不必要的知识丢掉。这种损失常常还未被注意,因为注意力集中在精细化的理论及其含义上,而不是集中在已被置于一旁的思想和要素上。显然这似乎就是效率这个中心概念的命运。不知怎么，效率被改换为只是配置效率。事实上，人们相信经济学的巨大成就之一是莱昂内尔·罗宾斯(Lionel Robbins)在20世纪30年代出版的著作，在这部著作里，经济学被定义为稀缺资源在可供选择用途中

进行配置的逻辑。市场要以它们对社会经济资源在可供选择的用途间作有效配置的观点来进行研究。

经济学的进步中疏漏了经营者和工程人员的效率观念，效率表明人员和机器的工作有多么好，或者有多么糟。一旦配置效率与效用或利润的极大化假设相结合，就不再有经营者和工程人员的效率观念存在的余地了。于是，企业和企业内部有可能进行次最优(suboptimal)经营的观念消失了。我相信，我1966年的那篇文章有助于恢复这个观念，即各种不同形式的次最优行为是可能的。这是该论文引起争论的一个原因。经营者、工程人员和心理学家都知道次最优行为，但是不知何故标准的经济理论总不乐意痛快承认次最优经营的可能性。X效率理论提出了标准理论在一个十分重要的方面不够完善的可能性。正是这种认为标准理论可能不完善的看法引起了强烈的反响。

X效率观念提出了基本的理论问题。什么是使标准理论局限于不完善的限制前提？如果次最优行为确实存在，我们应该怎样重新构造基础理论，尤其是厂商理论，才能容易和直接地对它作出解释呢？

X效率理论也提出了以经验为根据的问题。第一个以经验为根据的问题是，X低效率事实上是否重要。第二个以经验为根据的问题是，重新构造的厂商理论的内容是否会适合已知事实与尚未进行的经验研究。弗朗茨关于经验研究的三章论述，相当有力地指出，对上边这两个以经验为根据的问题都可以作肯定的回答。X低效率事实上显然很重要，而重新构造的厂商理论的内容，就它们通过了多次研究的检查而论，看来是得到支持的。

争论的另一个原因出于每种职业中都固有的那种基本的保守主义。X效率方法要求用另一种假设取代极大化的假设。这另一种假设我们可以把它称作极大非极大化假设。它允许非极大化行为，但是，当然也不排除极大化行为。注意到下列情况是重要的：与数学中一样，在一切科学领域中，为了得到另一种可供选择的意义和结果，改变基本假设是很正常的事情。说明这个道理的一个例子发生在1840年到1870年间的几何学领域，那时诞生了非欧几里得几何学。很清楚，这个道理对于物理学中的相对论和量子理论革命来说也是正确的，其他领域的例子也容易找到。这样的变化通常都伴随着争论，因此，当经济学试图改变其基本理论假设的时候，出现争论也就不足为奇了。

在我看来，可供选择的理论或模式应该以这样的方式进行阐述，即以经验为根据的结果要能够在可供选择的各种方法中作出决定。弗朗茨书中论述经验的三章，对这个问题谈得很多。我认为，一般说来，从事经验研究的作者通常总试图对正在考虑中的可供选择的方法作出合理的解释。

在一个问题上，我已明确地改变了想法，放弃我早期著作中的主张。我不再相信：惰性区域必须或通常由变动带来的利益和成本的平衡计算来决定。许多处于惰性区域的人可能不会想到改变的成本或利益。他们可能根本不进行任何计算。这是由如下事实必然得出的结论：各种各样的决定（如根据习惯或常规作出的那些决定）通常属于非计算类型。现在我的观点是，惰性区域是由人本性的基本心理和生理方面的状况决定的。换言之，如果我们必须对外部不管如何细小的刺激都作出反应的话，生活就几乎不可能继续下去。实际上，使我们集

中注意力的能力的背面，是我们不理会同时撞击我们的各种各样其他刺激的能力。这个情况也可以用“唤醒理论”来解释。要我们考虑采取任何种类的行动时，常常必须把我们唤醒，使我们处于超过某个临界水平的清醒程度才行。

强调本书很大程度上是罗杰·弗朗茨独自完成的，这点很重要，它针对X效率理论引起的问题提出了一种方法，X效率理论虽然博得明显赞成，但它仍然代表一种独立的判断。当然，我对每一方面的看法与弗朗茨不尽相同。很清楚正是由于弗朗茨的书对理论和事实两者提出了独立的解释，并试图小心地把两者联系起来，才使本书具有特殊的价值。

哈维·莱宾斯坦

于哈佛大学



前言 1

序言 1

1 导论 1

1.1 导言 …………………………………… 1

1.2 X效率和配置效率，以及由垄断力量引起的福利损失 ………………………………… 2

1.3 X效率和新古典的生产与成本函数 ………… 3

1.4 X效率理论的发展 ………………………… 4

1.5 X效率理论和一般化的新古典理论 ………… 5

1.6 作为一种研究计划的X效率理论 …………… 6

1.7 X效率理论的经验研究 ……………………… 8

1.8 对X效率理论的批评……………………………9

2 生产、成本和福利：述评 10

2.1 导言……………………………………………10
2.2 厂商的短期行为 ………………………………12
2.3 厂商的长期行为 ………………………………14
2.4 厂商、市场和效率…………………………… 18
2.5 各种隐含的意义 ………………………………23
附录：管制理论和X效率…………………………24

3 X效率：学术背景和理论介绍 39

3.1 导言……………………………………………39

3.2 复合的目标函数……………………………40
3.3 X效率理论：简介…………………………50
3.4 结论……………………………………66

4 X效率理论1 67

4.1 导言……………………………………67
4.2 X 效率理论中的个人 ……………………70
4.3 个人努力和惰性区域 …………………… 83
4.4 结论……………………………………91

5 X效率理论2 92

5.1 导言……………………………………92

5.2 个人和集体努力的企业内部的决定因素 …… 92
5.3 生产率、努力协定和囚徒的两难困境……… 99
5.4 市场结构、压力和努力……………… 105
5.5 综合和实例……………………… 109
5.6 含义和结论……………………… 115
附录：管制与X效率……………………116

6 经验证据：受管制的行业 130

6.1 导言……………………………… 130
6.2 电力公用事业…………………… 131
6.3 地方政府服务…………………… 151
6.4 交响乐团 ……………………… 153

6.5 航空公司……154
6.6 结论……156

7 经验证据：所有权形式 157

7.1 导言……157
7.2 所有者控制与经理控制的企业……158
7.3 公有和私有……180
7.4 结论……186

8 经验证据：市场结构 187

8.1 导言……187
8.2 产出和投入比率……187

8.3 利润和X效率 …………………………… 200
8.4 固定价格行为 …………………………… 210
8.5 结论 …………………………… 212

9 X效率：它的批评者和答复 213

9.1 导言 …………………………… 213
9.2 寻租 …………………………… 214
9.3 闲暇作为产出 …………………………… 218
9.4 竞争下的管理效用 …………………………… 227
9.5 产权 …………………………… 228
9.6 一些总的评论 …………………………… 234

10 意义和结论 236

10.1 X效率和配置效率 …………………… 236

10.2 作为一项研究计划的X效率理论 ……… 238

10.3 X效率及其批评者 …………………… 242

10.4 X效率理论和新古典主义理论：一些最终想法 ………………………………… 243

导论

1.1 导言

几年前，在普林斯顿大学召开的讨论行为经济学的会议上，我生平第一次发表了一篇很短的演讲，评论我所收集的有关以经验为根据研究X效率理论的文献。也许由于我的言辞简短，听众反应彬彬有礼。但是稍后，有一个人对我提出两点评论。第一，他发觉报告使他感到兴趣，但承认他没意识到X效率在多大范围内存在。第二，他对这批文献中报道的研究结果有另一种——非X效率——解释。他在我们能够集中的极短时间内匆匆作出的解释(好在他敏于表达)与新古典理论完全一致。我对自己的回答是："的确如此。"

我提到这件事，因为在写作关于X效率理论这本书时，我每每自问，怎样能够使书中的特定章节，写得别人无法用另一种——新古典的——方法进行解释。最终我懂得了，要按这样写书，最好的办法或者是永远不要写成这本书，或者是永远不开始写这本书。问题不在于这样的潜在批评，因为在这种"公事"中，我们每人提出的各不相同的观点，是我们增

进知识的一种方式。而且在撰写关于X效率理论的著述时，十分明显，你提出的任何主张，很快并且常常会遭受另一种解释的反驳。我们专业训练中的新古典微观经济学部分的影响是如此强烈，以致这样的反应几乎可以说是本能的反射。在新古典微观经济学理论方面训练有素的大学生常常也有同样的反应，只是有一点不同，因为他们依靠我们评分，所以他们常常觉得必须有礼貌，因此提出问题时总是说，"它是不是可能……？"（他们瞒不了我，因为我知道他们真正想的是什么！）

因此，我毫不后悔我开始写作此书，并决心完成它。因此，我将以这个理论应该有的面貌，也就是根据它的本来面目来描绘它。此外，书中几乎全部的经验证据，不但符合X效率理论的含义，而且都是由我独自照实叙述的。

1.2 X效率和配置效率，以及由垄断力量引起的福利损失

在（微观）经济学中是否需要像X效率这样的概念呢？在X（低）效率一词中，X代表来源不明的非配置（低）效率。（1966年莱宾斯坦首次提出"X效率"这个名词时，其来源比今天还要不明。）1954年，哈伯格(Harberger)估计，由于垄断力量造成的福利损失约占美国国民生产总值(GNP)的千分之一。（在本书中，我称这种形式的低效率为"市场配置低效率"，这种类型的低效率与X低效率、价格配置低效率和隐含的价格配置低效率不同。）随后出现的其他估计也报道垄断力量和关税产生同样小量的福利损失。恰如莱宾斯坦所指出的，这些估计只考虑到垄断和/或关税引起的（净）产量和价格扭曲。虽然

某些产量和价格可能非常扭曲，然而总的来说扭曲不可能很大。根据哈伯格和其他人的估计，它们并不大。

莱宾斯坦还指出，这些估计假设厂商是成本极小化的，也就是说，它们“有效地”购买和利用了全部投入。而X效率理论的主旨，就是要说明，免受竞争压力的保护不但会产生市场配置低效率，而且还会产生另外一种类型的低效率：免受竞争压力的厂商明显存在超额的单位生产成本。因为这种类型的低效率的性质当时尚不明了，所以称作X低效率。

1.3 X效率和新古典的生产与成本函数

认为(微观)经济学需要X效率概念的另一个理由，在于新古典派理论假设厂商是根据它们的生产函数和成本函数进行生产的。换言之，假设厂商会在既定的投入和技术水平下，实现产量极大化，因此它们的单位成本是极小化的。如果实际情况就是这样，那么教科书中关于生产和成本的技术性质的描述就是精确的。因为假设在厂商既定的投入集合约束下，产量是极大化的，所以，一旦知道了投入和技术的数量和性质，也就知道了产出的水平。因而投入-产出关系是技术关系，也就是说，在投入集合和现行技术既定的情况下，实质上，是技术而不是决策本身决定产出。同样，因为假设对于任何既定的产出率(或产出变化)成本都是极小化的，所以成本-产出关系也是一种技术关系。

但是，如果生产和成本过程主要不是技术关系，那末教科书上的叙述就需要加以修改。X效率理论就是这样的一种修正：X效率理论假设，新古典理论仅在厂商(事实上)根据

它们的生产和成本函数进行生产的时候才是精确的。因此，X效率理论认为新古典理论是一种"适用范围有限的理论"。只有在厂商事实上确是根据它们的生产和成本函数进行生产的时候，新古典理论才最为适用。

在本书第1章，我将集中论述教科书上有关微观经济理论的技术和运行机制的性质，目的是为了说明，这个理论把出现X(低)效率的可能性假设掉了。反对X效率理论的新古典论点是多余的：在新古典的微观经济理论中，根本就没有X(低)效率的容身之处。

1.4 X效率理论的发展

X效率理论在20世纪60年代早期的发展，不是经济学家解释厂商行为的唯一尝试，对这种厂商行为，利润极大化、成本极小化模型似乎是无法解释的。鲍莫尔(Baumol，1959年)提出了一个以销售最大化为目标的厂商模型。鲍莫尔解释说，构造这个模型(至少是部分)的动力来自他在企业中当顾问的经验。此外，马里斯(Marris，1963—1964年)提出了厂商以增长为目标的模型，而威廉森(Williamson，1964年)的企业模型中的经理则对工资薪金感兴趣。

如同我已经提到的那样，莱宾斯坦提出X效率的概念，主要是受到(各种类型)事实数据的推动。其中一类事实是他观察到他在伯克利大学的研究生研究助手未能得到充分利用！正是这一事实使他想要知道，投入物未被充分利用是否是比这一例子更加普遍的现象。

另一个观察到的事实是正规的经济发展模型和他自己以

及他在欠发达国家工作的几个同事的经验之间存在差异。最后，可获得的事实似乎与新古典理论不符：（1）事实表明只要将工厂的内部组织作相当简单的变动，厂商就能够增加它们的产量；（2）事实表明，厂商并不是按照边际分析的原理经营的；（3）宏观数据表明，体力劳动和实物资本以外的某种东西在几个工业化国家的增长率中发挥着重要的作用。

所有这三种类型的事实都表明有必要对各种可供选择的厂商理论进行一次考察，以便能弥补只依赖市场配置效率这个概念的经济理论中出现的空白。怎样才能解释这三种类型的事实呢？莱宾斯坦提出的X效率理论就是对这些事实的解答。在第3章中，事实的表述与新古典理论和X效率理论的最初陈述似乎不相一致。此外，为使读者了解X效率理论形成的环境，对其他几个（非利润极大化的）厂商理论也都作了叙述。

1.5 X效率理论和一般化的新古典理论

X效率理论的性质，以及它与当前占支配地位的新古典理论在某些根本方面存在差异这一事实，意味着在两者之间进行比较是不可避免的事情。实际上，我已经好几次提到这一点。但是，现在据说新古典理论有两种形式：（1）假设利润极大化和成本极小化的形式；（2）假设效用极大化和成本极小化的"一般化"形式。而在这种一般化的形式中，成本是在雇员选择使他们的效用极大化的行为方式约束下实现"极小化"的。举例来说，如果假定雇员是效用极大化者（包括在工作中偷闲的选择在内），那末，厂商的单位成本极小化就受到雇员目标的约束，这个目标就是不以脑力或体力的最大努力进行

工作。

X效率理论的某些假设"符合"新古典理论的这个一般化形式。也就是说后者采用了能够"解释"上述第三种类型事实的那些概念，即X效率理论的核心假设，以及与X效率理论的含义相一致的经验结果。例如，X低效率可能产生于这样一个事实，即劳动合同是"敞开的"，这使工人可以随意决定自己的努力程度。又如，"惰性区域"假设，指的是个人想找到一个舒适的努力水平并倾向于保持这个努力水平，有时似乎不顾及环境的要求。倘若存在一种合适的交易水平和控制成本，那么就可以争辩说，X效率理论的核心假设与新古典理论的一般化形式是一致的。在第9章里我将论述一般化的新古典理论是怎样对X效率理论的某些经验研究进行"解释"的。

1.6 作为一种研究计划的X效率理论

X效率理论与一般化的新古典理论不相一致，是一个十分重要的方面。X效率理论假设，个人可以是效用极大化者，也可以不是效用极大化者。更确切地说，不是所有环境都导致个人表现出最大化——经仔细计算——的行为。换言之，极大化是可以观察到的，而不是假设出来的。假设极大化行为，就是假设不存在X低效率。这就是假设为何重要的一个原因！

这一点差异使X效率理论在一些重要方面与新古典理论相去甚远。表1-1是用来说明这个问题的①。存在两种关于

① 这种描述X效率理论和新古典理论的方法是我的同事哈林德·辛格提出的。在罗森(Rozen)的著作中(1985年)也有论述。

个人行为的假设：（1）个人（总）是效用极大化者，和（2）个人并不（总）是效用极大化者。同样也存在两种个人采用的实际行为类型：它们或许能极大化，或许不能极大化。因此，有可能造成两种类型的"误差"：类型1和类型2。在这里，类型1误差的意思是：假设个人不想实现极大化，可事实上他却实现了极大化。类型2误差的意思是：假设个人想要实现极大化，可事实上他没有实现极大化。由于新古典理论忽略非极大化行为的可能性，因此通过假设，它把类型1误差确定为零。

表 1-1　隐含的类型1和类型2误差的方法论结论

实际的 / 假设的	极大化	非极大化
极大化	无误差	类型2误差
非极大化	类型1误差	无误差

相应地，新古典理论框架使得类型2误差出现的概率变得很多。事实上，新古典理论正是在这种情况下研究类型2误差的。这就是说，存在着一种有意识的努力以说明那种乍一看来是非极大化的行为，在考虑到所有相关因素以后，事实上都是极大化行为。犹如我将在第1章和第8章指出的那样，这些因素通常被视为"丢失的环节"。现在，我先简单地给出这些因素中的几个名称：效用、寻租、风险规避和产权。

与新古典框架恰成对照，X效率理论既承认极大化行为的可能性，也承认非极大化行为的可能性，因此，它既承认

类型1误差的可能性，又承认类型2误差的可能性。作为一种研究计划，X效率理论能更好地认识这两种误差间的权衡，并尝试把两者保持在可以处理的限度之内，因而经验检验能够实现。从表1-1的对角项上也可以看得出：当个人采取极大化行为时，他是被假设采取极大化行为的；当个人没有采取极大化行为的时候，他是被假设没有采取极大化行为的。故而，由于新古典理论假定所有行为都是极大化行为，所以它就把自己限制在类型2误差（象限4）内。但是，在这方面新古典的分析试图说明，为什么所有行为（极大化的和“表面”非极大化的）实际上都归入象限1：个人总是采取极大化行为者。

另一方面，X效率理论代表一种更为全面的研究计划，因为它允许行为可以与四个象限中的任何一个保持一致。通过承认这个可能性，X效率理论穷尽了各种可能发生的事情。本书的第4和第5章将用以介绍这个理论。第4章集中论述单个工人的心理，第5章考察作为集团成员的个人，并讨论管制以及它对配置效率和X（低）效率可能产生的影响，和为减轻两种形式的低效率而制订的某些政策。在这两章中，我们将考虑表1-1所示的理论，以及X效率理论的研究计划与新古典理论的研究计划有何不同。

1.7 X效率理论的经验研究

第6章到第8章讨论与X效率理论意义相一致的经验研究。每章论述一个论题范围的研究。（这几章中报道的某些研究，似乎与论题不太“符合”，原因有二：（1）某些研究

已越出了一个论题的范围；（2）我不知道应该将它放在别的什么地方！）第6章描述管制企业特别是公用事业的经验证据。第7章描述所有制形式，包括公有制和私有制的证据。第8章描述有关市场结构的证据。

1.8 对X效率理论的批评

第9章论述反对X效率理论的一些论点和某些与它的含义一致的经验研究。将对以下四个"学派"的批评意见进行讨论：（1）强调寻租的学派（第1章和第4章的附录中也对此作了讨论）；（2）强调在职闲暇作为产出的学派；（3）强调在追求效用极大化的经理中间存在竞争的学派；（4）强调财产权的学派。尽管事实上这些学派在某些方面是彼此重叠的，但这样的分类证明还是有用的。如我所要论证的，对X效率理论的这些批评，使用的都是已被广泛使用而不是独到的词语。对于贯穿本书的X效率理论和新古典理论，我的结论是，相信它们仍有优点和弱点。最后，我还要为今后研究X效率理论提出一些建议。



生产、成本和福利：述评

2.1 导言

经济学关注资源的有效使用。通常，这指的是通过价格和市场进行资源配置的意思，因而这也就是"市场配置(低)效率"这一术语和市场如何在各种竞争性用途间配置资源这一问题的意思。从市场有效配置资源的能力——为每一种商品与劳务的生产配置适当数量的资源——的意义上讲，厂商能够影响市场配置效率，这种影响的大小，取决于厂商或组成行业的厂商集团所掌握的市场力量。

同任何理论一样，厂商理论也提出某些假设。一个假设是，厂商是内部有效率的。就是说，在既定的投入下，厂商生产极大的产出(技术效率)，因此厂商是成本极小化者。这样说是否符合实际，假设的现实性是否重要，这里无须争论。重要的是这些假设使得经济理论可以忽略厂商的内部运行和厂商的内部效率，而集中注意市场效率。本书的目的在于说明，存在着一种类型的(低)效率，过去，由于存在这些假设，所以它一直被经济理论所忽视。而且，这种效率可能比配置效

率更加重要，它的存在和规模将被证明具有重要的政策意义。哈佛大学经济学家哈维·莱宾斯坦把这种类型的效率称为X效率。

为了弄懂X效率，有必要先理解市场配置效率和绝大多数标准微观经济学理论的教科书中阐述的传统厂商理论。读者应该明白，本书假设的微观经济学理论是以中级微观经济学教科书的内容为代表的①。本章的目的，是对微观经济学中关于生产、成本和福利理论的某些方面作一个评论。本书讨论的内容以这个理论作为基础，不仅是X效率理论，而且还有其他一些理论都将在这个基础上展开讨论。因此，我们假设读者已经上过一门经济学课程，讨论过等产量-等成本分析和垄断力量对福利的影响。没有完成这样一门课程的读者，可能会觉得这方面的材料有些简略。因此，我们建议这些读者能仔细地阅读一本中级微观经济学教科书，以便熟悉生产理论、成本理论和垄断力量的福利效应。

把微观经济学理论描述为一种忽视非配置效率的理论，是因为它对短期乃至长期生产与成本理论的阐述隐含了内在效率的假设。这将使读者有机会懂得，技术效率和成本极小化的假设如何互相关联，它们合在一起，又如何导致非配置效率被排除在厂商理论之外。下节将阐述厂商的短期行为，再下节讨论厂商的长期行为。本章最后讨论市场配置效率和垄断力量的福利成本。

① 在中级微观经济学水平上，海因茨·科勒（Heinz Kohler）1986年出版的教科书对X效率理论作了相当全面的论述。在初级微观经济学水平上，最详尽论述X效率理论的是希勒（Schiller）的著作（1986年）。

2.2 厂商的短期行为

2.2.1 短期生产函数

在短期，至少有一种投入是固定的，在此期间，投入与产出的关系表示为Q=f(K,L)，这里Q表示产出，K表示资本，L表示劳动。这种关系称短期生产函数，要记住的重点是，它指出了在资本存量不变的情况下，任何一定量劳动可以获得的最大产出。这种情形如图2-1所示。从图中可以看出，1单位劳动同1单位资本相结合，能够获得10单位的产出；而2单位劳动与1单位资本结合则能获得30单位的产出，如此等等。此外，对于每个资本量都可以估计出一个不同的短期生产函数。这在图2-1中也看得出来，如$Q=(\bar{K}_1,L)$，$Q=f(\bar{K}_2,L)$，其中$K_1<K_2$。因此，图2-1表明，较大的资本量导致每单位劳动的较高产出。尽管每单位劳动的

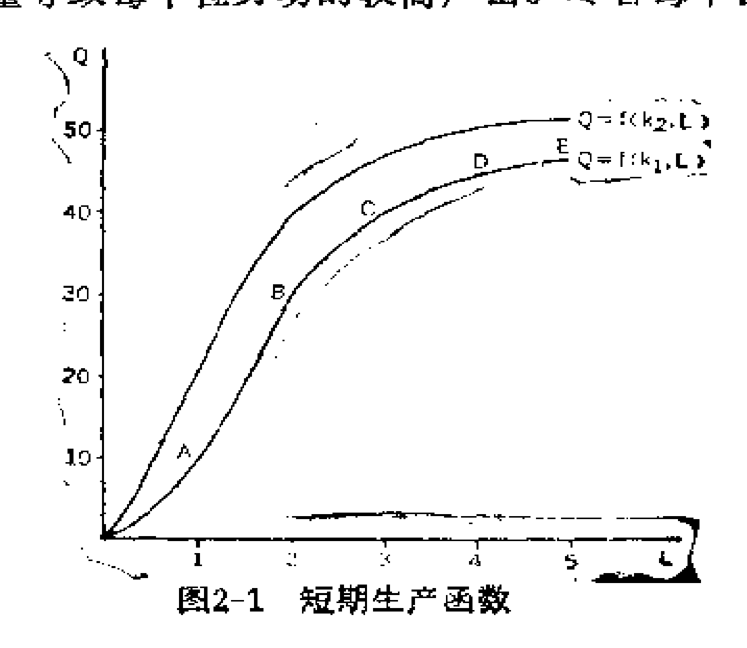


图2-1 短期生产函数

产出随着资本/劳动比率的增加而增加，但我们仍然必须记住，每个短期生产函数表明了与一个固定资本量相结合的每单位劳动量的最大产出率。

因为生产函数指的是由一定量劳动可以获得的最大产出率，所以，它也是指为获得一定产出率所必需的最小劳动量。因此，我们从图 2-1 中也可以看出，假设仍旧使用 1 个单位资本，10个单位的产出最少可由 1 个工人来生产，而30个单位的产出，最少需要有 2 个工人生产，如此等等。

2.2.2 短期成本函数

很明显，厂商的支出来自生产过程。因此可以认为，它们的生产函数和成本函数是相关的。再来考虑表示下述劳动-产出组合的生产函数 $Q=f(K_1,L)$：$L=1,Q=10$；$L=2,Q=30$；$L=3$，$Q=40$；$L=4,Q=45$；$L=5,Q=46$。图 2-1 中，这些关系分别用 A 至 E 点表示。为使读者懂得如何把这些点画在厂商的成本曲线上，我们假设劳动价格(P_l)为20美元，资本价格(P_k)为 20 美元。那末这些产出率中每一种情况的总成本 TC(这里总成本是根据公式$TC=P_lL+P_kK$ 计算的)就是：$Q=10$，$TC=40$美元；$Q=30$，$TC=60$美元；$Q=40$，$TC=80$美元；$Q=45$，$TC=100$美元；$Q=46$，$TC=120$美元。

这些组合就是厂商的短期总成本曲线。而且，因为$ATC=TC/Q$，所以厂商的平均总成本曲线可以推算出来。在图2-2中，这 5 个产出-平均成本组合分别以 A 至 E 各点表示。因为生产函数指的是既定投入(成本)的最大产出，所以它也暗指获得一定产出率所必需的最小投入(最小成本)。后者就是

厂商的成本函数：获得一定产出率在技术上所必需的最小成本水平。虽然成本隐含在生产函数里，但是一旦我们知道投入价格，我们就可由这样的成本函数估量最小成本。很清楚，这两种函数是互补的。例如，图 2－1 中A点指的是使用 1 个单位资本，1 个工人能（假定可能获得的最大产出就是实际产出率）获得 10 个单位产出。图 2－2 中A点是指 10 个单位的最小平均成本是 4 美元。这时，更低的单位成本是不可能的，而更高的单位成本却是可能的。

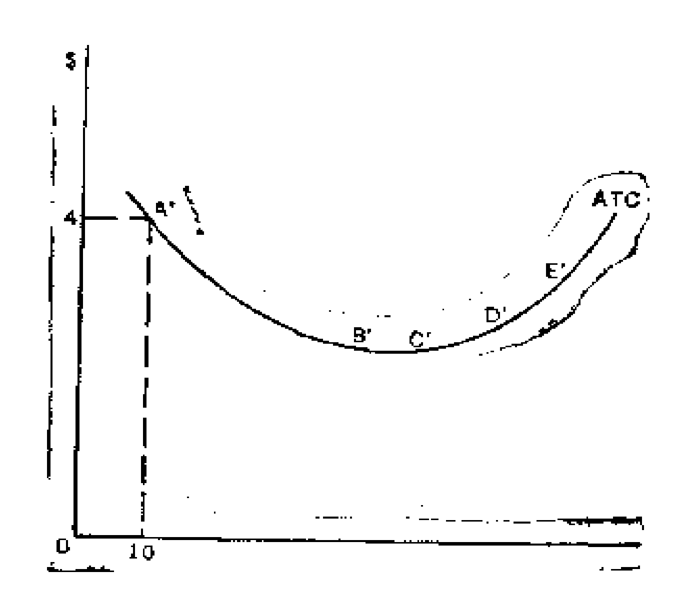


图2-2 短期平均总成本曲线

2.3 厂商的长期行为

2.3.1 长期生产函数

在长期内，即在足以使所有投入量都成为变量的时期内，任何产出率都能够通过众多潜在资本/劳动组合中的一

个来获得。能够达到产出率Q_1的组合叫做等产量线。在图2－3中，我们把产出率Q_1表示为一条等产量线，它的斜率等于$-(MP_l/MP_k)$，其中MP_l和MP_k分别代表劳动和资本的边际产量。也就是说，为了使产出保持不变，各种投入可以按等于它们相对边际产量的比率彼此替代。因此，如果$MP_l=2MP_k$，则厂商每增加1个单位劳动力就能减少2个单位的资本存量。

在图2-3中，等产量线Q_1表明Q_1可以用比如5K和1L获得。在图2-3中，我们画出一系列等产量线，每一条等产量线代表不同的产出水平，而且$Q_1<Q_2<Q_3$。

关于等产量线的重点是，它指的是获得一定产出率所必需的最小投入量。例如，在图2-3中，如果Q_1用5个单位的资本生产，那么必需的最小劳动量就是1个单位。因此，我们也可以把等产量线Q_1解释为，如果使用1个单位劳动，那末必需的最小资本量就是5个单位。

因为等产量线表示获得一定产出率所必需的最小投入

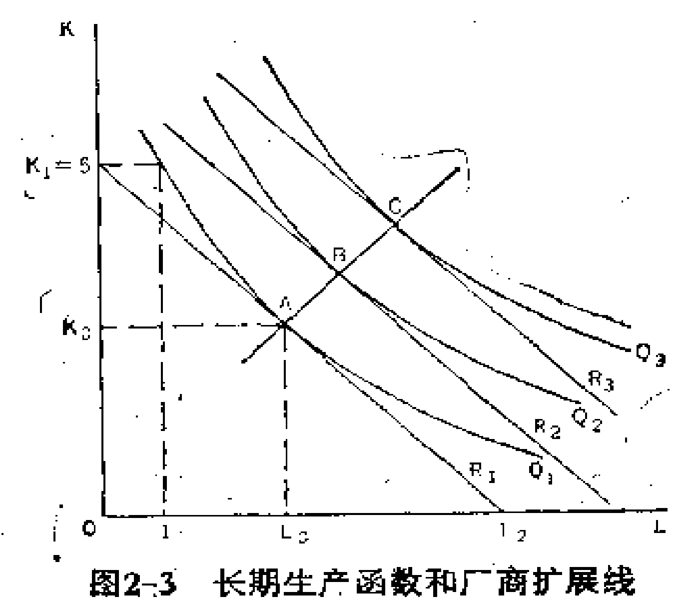


图2-3 长期生产函数和厂商扩展线

量，所以它也表示由一定投入可获得的最大产出率。因此，图2-3也表示Q_1是可以由这个等产量线所代表的任何资本-劳动组合获得的最大产出率。一定产出率的最小投入或一定投入的最大产出具有相同的解释，即每一个都是另一个的对偶。

2.3.2 最优——成本极小——投入组合

对于既定的产出率，为使生产成本减到极小，厂商能够采用何种投入组合呢？原则上，任何由等产量线上的一个点所代表的投入组合都能使厂商的成本实现极小化。但是，事实上，每一组投入价格都只允许一个投入组合使厂商实现成本极小。为了确定这个使成本极小化的投入比率，厂商需要知道几件事情。第一，厂商必须知道既定的支出率（R）能够购买多少投入。购买投入的能力取决于R，劳动的价格（P_l）和资本的价格（P_k）。因此，$R=P_lL+P_kK$。重新整理一下，我们可以得到$K=R/P_k-(P_l/P_k)L$，$L=R/P_l-(P_k/P_l)K$。既定R可以购买的所有资本-劳动组合称作等成本曲线。因此，如果$R=100$美元，$P_l=20$美元，$P_k=20$美元，那末$K=5-L$，$L=5-K$。从而，如果$K=4$，那末$L=1$；如果$K=3$，那末$L=2$，以此类推。然后，既定等成本曲线上每一点代表的支出或成本水平是相等的。图2-3表示出斜率为$-(P_l/P_k)$的等成本曲线R_1。要点在于，等成本曲线表示的是既定R能够购买的最大投入量。回想一下对偶概念，我们也能够说，等成本曲线表示购买一定量由等成本曲线给出的资本和劳动所需要的最小R。图2-3也画出一系列等成本曲线，其中$R_1<R_2<R_3$。

为了实现成本极小，厂商要做的第二件事是，选择一种使

得花在每种投入上的每一元钱的边际产量相等的资本-劳动比率。也就是说，厂商使成本极小的投入组合就是能使$MP_l/P_l = MP_k/P_k$的那个投入组合。经整理，我们得$MP_l/MP_k = P_l/P_k$。后一个表达式如图2-3的A点所示。等产量线(Q_1)和等成本线(R_1)的切点——在这点上，两条线的斜率相等——就是对于产出Q_1的成本极小的投入组合。同样，我们可以说，B点是Q_2的成本极小的投入组合，C点是Q_3的成本极小的投入组合。这些切点(A、B、C)连起来称作扩展线。因此扩展线上的每一点指的都是在长期内获得每个产出率的最小成本。扩展线上的每一点就这样描绘短期和长期成本函数。

2.3.3 扩展线与长期成本函数

为了更好地理解扩展线怎样找出长期成本曲线，请看图2-3。在支出水平R_1下，厂商能够购买OL_2或OK_1。因为OL_2和OK_1都在同一等成本曲线上，所以它们代表相同的支出率。同

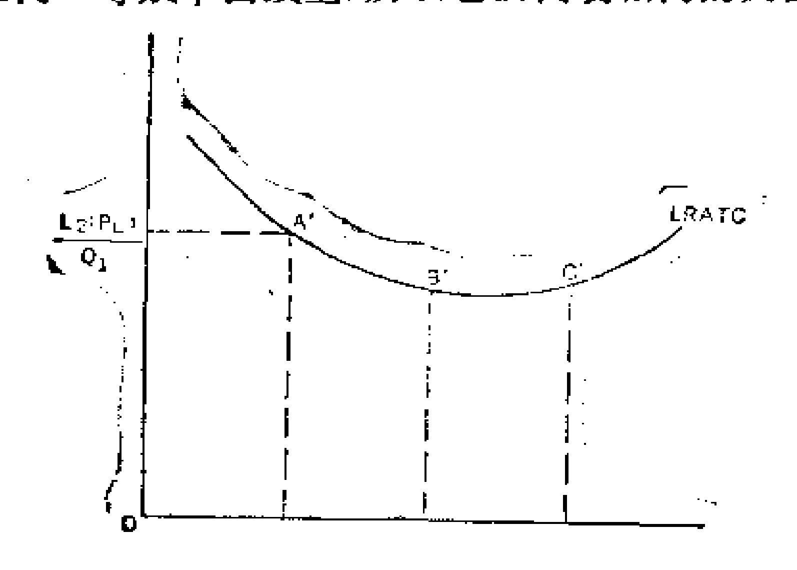


图2-4 长期平均总成本曲线

样理由，它们每一个都等于购买最优投入组合（A点表示）所要求的支出率。因此，产出Q_1的总成本就是$OL_2(P_l)$或$OK_1(P_k)$或$OL_0(P_l)+OK_0(P_k)$。因而，平均总成本就是$OL_2(P_l)/Q_1$。在图2-4中，这由长期总成本曲线上的A′点表示。同样，点B′和C′分别代表产出率Q_2和Q_3的最小单位成本。要理解的重点在于，每一点因此都代表获得既定产出率的最小长期成本。

2.4 厂商、市场和效率

2.4.1 导言

这里描述的生产函数和成本函数都假定厂商在投入水平既定的情况下使它的产出达到极大化(生产函数)，以及在产出水平既定的情况下使它的成本达到极小化（成本函数）。在这些条件下，可以把厂商视为一个"黑箱"，它只是以技术上必需的最小成本把（既定）投入转变为(最大)产出的技术关系。如果我们假定厂商按它的生产函数和成本函数营运，厂商内部发生些什么就变得无关紧要。如果厂商真的能根据既定投入实现产出极大化以及为获得既定的产出率实现成本极小化，那末就几乎没有什么别的东西需要说的了。不管厂商内部有什么具体活动发生，厂商都被假定是有效率的。因此，它们"不再引起人们的兴趣"。这就好像厂商理论没有包含厂商一样。

2.4.2 市场配置效率

一旦断定厂商内部有效率——厂商按它的生产函数和成

本函数生产——经济学家对效率的关注就集中到市场上。也就是说，市场能有效地配置资源吗？为了理解这个问题的答案，我们要使用图 2-5 与图 2-6 。图 2-5 以 P_e 和Q_e 表示市场均衡，就是 $Q_d=Q_s$ 的价格和数量。我们还表明，需求曲线反映消费的边际收益(MB)，供应曲线反映生产的边际成本(MC)。因此，市场均衡就是使MB＝MC的价格-产出组合；对于这个组合来说，总收益减去总成本的差因而也就达到了它的极大值。因此，市场均衡就是能使净经济福利极大化的价格和数量水平。由此可以推得，任何小于 Q_e 的产出率就是MB (等于 AQ_1)超过MC(等于 BQ_1)的产出率，因此，净经济福利将随着产量增加而继续增加。另一方面，任何大于Q_e 的产出率就是 MB(等于CQ_2)小于MC(等于DQ_2)的产出率，因此，净经济福利将随产量增加而减少。

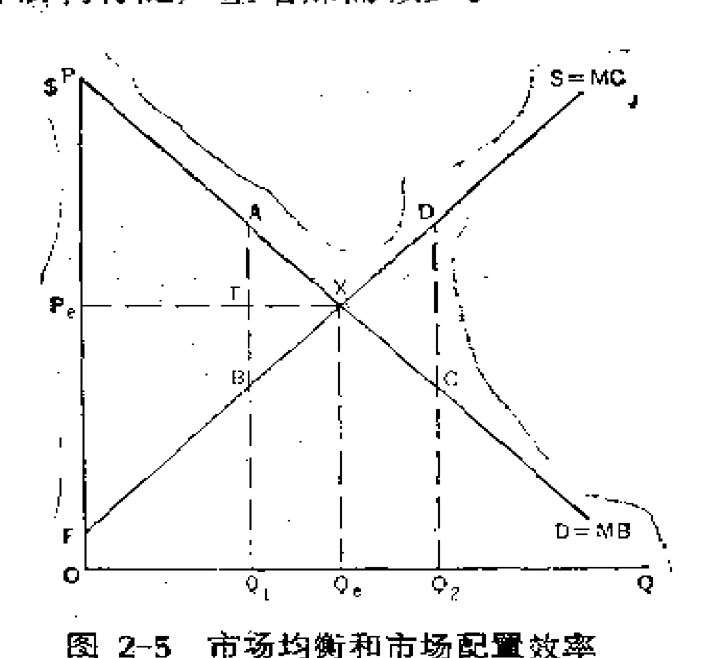


图 2-5　市场均衡和市场配置效率

我们还能看到，产出率Q_e是使P＝MC的产出率。这个条件十分重要，因为它表明生产出来的商品的价值等于所使用

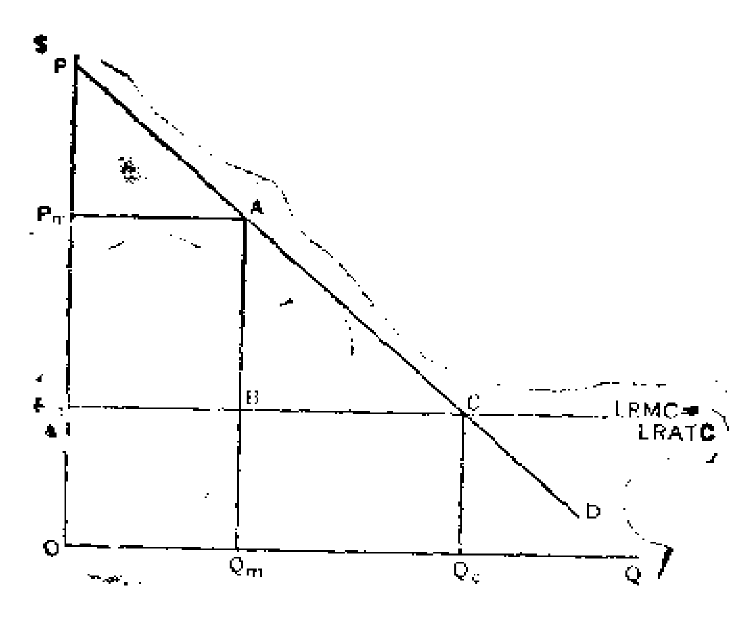


图 2-6 厂商和市场配置效率

资源的机会成本。也就是说，所使用资源的价值正好等于它们的机会成本。与此相对照，对小于Q_e的产出率来说，$p>MC$，这表明资源现在使用的价值超过它们在下一个最佳机会中使用的价值。换言之，低于Q_e的任何产出率，不是社会最优产出率：社会偏好更多的是这种商品(p)，而偏好相对较少的是另一种商品(MC)。因此，小于Q_e的任何产出率表示资源的不足配置，也就是说，市场没有把足够的资源配置给这种商品的生产。

与此相比，对任何大于Q_e的产出率来说，$p<MC$，可使用与$Q<Q_e$情况下同样的推理，大于Q_e的产出率表示资源的过度配置，也就是说，市场把过多的资源配置给这种产品的生产。大于Q_e的产出率也不是社会最优产出率。

可以证明，市场均衡产出率也表示福利的极大化，因为它使消费者剩余和生产者剩余的总和实现了极大化。消费者剩余以消费者对他们收到的东西所愿意支付的最大金额和它

们实际为此支付的金额（实际支出）之间的差额衡量。因此，消费者剩余是总收益和总支出的差额，即消费者的净收益。因为需求曲线反映边际收益，所以总收益可用需求曲线以下的区域来衡量。就是说，总收益只是每个消费单位收益（MB）的总和。在图 2-5 中，Q_e 的总收益是 $OPXQ_e$，总支出是 $OP_e \times Q_e$。因此，消费者剩余是（$OPXQ_e - OP_eXQ_e$），即 PXP_e。

生产者剩余是以厂商获得的总收入和他们以任何特定产出率供应时愿意接受的最小价格之间的差额来衡量。因为供应曲线反映边际成本，所以它也反映厂商生产额外一个单位产出愿意接受的最小价格。因此，就产出率Q_e为例来说，厂商要求得到的最小支出是 $OFXQ_e$，而实际收到的总收入是 OP_eXQ_e。因而生产者剩余等于（$OP_eXQ_e - OFXQ_e$）或P_eXF。

可以证明，任何其他产出率都不能使消费者剩余和生产者剩余的总额达到极大。例如，如果产出是Q_1而不是Q_e，那末消费者剩余减少面积AXC。这就是说，在Q_1和Q_e之间还存在这样一些消费者，他们愿意支付比均衡价格更高一些的价格。这些消费者不能实现这种剩余，是由于（如果你能记得的话）资源不足配置即配置低效率所致。在 Q_1，生产者剩余要减少EXB。这就是说，在Q_1和Q_e之间还存在这样的生产者，他们得到的价格比他们的最小价格要高。这些生产者因为资源配置不足而没有实现这部分剩余。

对高于均衡水平的产出率来说，边际收益会低于消费者不得不支付的价格，而生产者接受的最低价格将超过市场价格。因此，超过均衡水平的产出率是不合意的。所以只有均衡的产出率——有效率配置——能使消费者和生产者剩余的总和达到极大，当达到市场配置效率时，福利就实现了极大化。

图 2-6 也显示了市场配置效率。鉴于图 2-5 表示的是一般市场的配置低效率，根据市场均衡和不均衡的观点，图 2-6 从一个完全竞争的厂商和一个具有市场力量的厂商角度研究这个问题。这使我们还可以研究市场力量的福利成本这样一个更为广泛的问题。由于我们再次假定不管厂商是否是完全竞争的，它都是内部有效率的，于是经济学理论就集中研究市场对资源进行配置的效率。因此效率就与市场配置效率同义。

在图 2-6 中，竞争厂商以价格P_c生产产量Q_c，而垄断厂商在较高价格 P_m上生产较少的产量Q_m。不管市场结构如何，我们假定这个厂商是成本不变的产业的一部分。因此，不论产量多少，长期平均总成本(LRATC) 不变。因此，LRATC＝LRMC。

为了把市场配置效率作为市场力量的福利成本的一部分来讨论，我们从估计在完全竞争和存在市场力量情况下的消费者剩余数量开始。在完全竞争情况下，消费者的总收益或他们愿意为产出OQ_c支付的货币数量在图 2-6 中为$OPCQ_c$。另一方面，消费者将支付 $P_c(Q_c)$或OP_cCQ_c。因此，在完全竞争情况下消费者剩余是PCP_c。在垄断条件下，消费者为得到产出 OQ_m愿意支付$OPAQ_m$。而另一方面，他们付出的都是 $P_m(Q_m)$或 OP_mAQ_m。在垄断条件下，消费者剩余为PAP_m。因此市场力量使消费者剩余的数量减少了 (PCP_c-PAP_m) 或P_mCCP_c。完全竞争使消费者剩余达到最大，因为它导致相对低的价格和相对高的产出率，对于这个产出率来说，$p=MC$。

这块损失掉的消费者剩余分为两部分：(1) ABC，(2) P_mABP_c。而积P_mABP_c是垄断利润。对产出 Q_m，垄断厂商要

价P_m，但平均成本只有LRATC（$=P_c$）。面积P_mBBP_c因此等于利润$(P_m-ATC)Q_m$。但是利润可以看作是收入从买者到卖者的转移。买者放弃的东西恰好等于卖者得到的东西，所以没有损失。因此，这些利润不表示福利损失。

面积ABC表示市场配置低效率。与在P＝MC处生产Q_c的完全竞争厂商相比，垄断厂商要价较高而生产较少。而且，在产量Q_m处，价格（等于AQ_m）超过边际成本（等于BQ_m）。因此，在Q_m水平上，市场是配置低效率的：市场没有把足够的资源配置给这种商品的生产。这种市场配置低效率是完全的福利损失，因为消费者得不到他们所希望的商品组合：他们得到的商品，一种太多，而另一种却不够。

因此垄断力量的成本就是对比希望产出率低的产出率支付相对高的价格，以及对消费者所希望的商品组合的扭曲。再次地，这个分析还隐含有这样一个假设：不论是完全竞争厂商还是垄断厂商，他们都分别以同样的平均成本LRATC和边际成本LRMC进行生产。就垄断厂商比完全竞争厂商具有较高的生产成本而言，我们归之于利润的东西中的某些部分，实际上是较高的生产成本。这里的各种隐含意义，即本书和X效率理论的关键，将在随后几章中讨论。然而下一节将引入这些观念。

2.5 各种隐含的意义

在本章所述的厂商理论中，假定厂商都有一个以固有方式实现的单一目标。这个目标被假定是利润极大化。第二，在实现利润极大化过程中，厂商进一步被假定会对环境变化

能进行平滑迅速的调整。例如，劳动价格的相对提高预计会导致每单位产出的劳动/资本比率下降。或者，劳动边际生产率的（相对）提高，预示着每单位产出的劳动/资本比率提高。最终，假如其余情况不变，厂商用于支出的可获得基金的增加预示着产出的增加。第三，假定厂商能像教科书所说的那样经营，就是说，它使利润实现极大化，使成本实现极小化。对于既定的投入，他们使产出极大化；对于既定的产出率，他们使成本极小化。第四，假定厂商的行为与它的规模和所有制形式无关。对待厂商就好像对待单一的个人那样。下一章直至全书终了，我们将对这种观察厂商的框架提出质疑，并阐述另一种可供选择的框架——X效率理论。

附录：管制理论和X效率

A.2.1 管制、X效率和微观理论：一般考察

正统微观经济学理论与X（低）效率的存在基本不相容。本章我们叙述标准经济学教科书对于生产与成本的看法，有意强调它的僵硬机械部分和其成本极小化行为的假设。我们还指出，这种市场均衡的标准论述假定不存在净收益的浪费或机会的未充分利用。也就是说，当市场达到均衡时不可能存在任何X低效率。总之，我们试图表明，微观经济学理论与X效率的存在完全不相容。有关反对X效率的观点，人们想说出多少，就可以说出多少。然而，事实是只要他的研究没有跳出教科书中微观经济学的理论框架，那么他反对X效率理论的那些论点就都是多余的。这就是说，在正统教科

书的微观经济学框架里，从事研究的人们总能找到为什么X效率不存在的“有力的理论”根据。但是，这种结论实际上在模型的假设中就已经给出了。现在的若干政府管制理论，就是这方面的一个例子。

我们从管制是一种商品，从而可以在市场上进行交易这样一个假设开始。我们进一步假定所有个人都是有理性的。这样，市场应该是出清的，因为只有在市场均衡时，才不存在来自交易的收益。换言之，在市场均衡时，不存在尚未利用的(无需成本的)获取净收益的机会。也就是说，在市场均衡时，X低效率的程度等于零。因此，从管制就是有理性的个人在市场上进行交易的商品这样一个假设开始，由此假设不存在X低效率，从而也就是认为X低效率理论本身“无效”。在下面的论述中，请读者记住，我并不是想从这些作者或他们所阐述的理论那里去掉什么东西，而是想着重地说明现在的几种管制理论与X低效率的存在完全不相容。用这些理论中的任何一个作为“证据”来反对X效率理论的存在或重要性，在我看来都是没有意义的。因此，为了便于阐述，我将只提供这几种管制理论详情的轮廓。我将从这些理论的前身公共利益理论开始。

A.2.2 公共利益理论

管制的公共利益理论并不把管制看作商品，而是看作政府对一种公共需要的反应。这个理论① 似乎可以追溯到1877

---

① 包括本文所引案例在内的关于管制的公共利益理论发展简史，可参见艾尔弗雷德·卡恩(Alfred Kahn)1970年的著作。

年，当时法院在对芒恩诉伊利诺伊州(Munn v. Illinois)的判决中说，通过管制来保护客户免受私营垄断厂商侵害是正当的。另外两个法院的判决在这里也很有意义。第一个是，1932年在纽约艾斯公司诉莱伯曼(Ice Co.v.Leiberman)，法院判定：为了防止商务和(或)社会受到过分强烈的无管制竞争带来的损害，管制商务是合适的。第二个是，1934年内比亚诉纽约州(Nebbia v. New York)的案件，法院判定，如果从立法者的立场看，对任何行业实行的经济管制符合公共利益，那么，只要立法者不是以任性或歧视的态度办事，美国宪法就不阻止这种管制。而不论被管制的是公用事业还是其他行业都没有什么区别。

在对管制的公共利益理论（和别的理论）所作的研究中，波斯纳(于1974年)争辩说，这个理论或明或暗地包括着这样一个假设，即市场是脆弱的，如果放任自流就会趋向不公正和(或)低效率，而政府管制是对社会的公正和(或)效率需求所作的无代价、有效和仁慈的反应。

在它的“自然”垄断变体中，政府特许垄断形式的管制之所以符合公共利益，有着两个明显的理由：那就是取决于假定得自规模经济的收益和与规模经济无关的收益。就规模经济而论，自然垄断厂商是成本结构以由相当高的固定成本组成为特征的一种厂商。因此，如果市场上有两家或多家厂商的话，那么，同它们相比，自然垄断厂商的成本乃至价格将要低一些。此外，自然垄断的业务性质(如公用事业)使竞争不再可能存在。就不取决于规模经济的论点而言，据说自然垄断厂商供应的是必需品，而且，人们有时认为，由于竞争造成的设施(电话线)重复会给客户带来不便。

人们认为，管制符合公共利益还因为需要保护稀缺资源。例如，有限的收音机波段数目导致了对电波波长的管制。

但是，报酬率形式的管制已被证明会造成两种类型的低效率：配置低效率和X低效率。后面我们还要谈到这两种低效率。现在我们的兴趣只在于指出，新古典传统的另外几种管制理论，是如何将X(低)效率简单地假设掉的。

A.2.3 施蒂格勒模型

施蒂格勒(Stigler, 1971年)提出的理论与管制的公共利益理论形成鲜明的对照。他运用许多政府机构的例子，如民用航空委员会对进入航空业的影响，联邦存款保险公司对限制商业银行进入的影响，以及证券交易委员会限制小型互助基金会(因而就减少了大型基金会的销售成本)的影响，他得出结论说，经济管制主要不是政府对公共需要的有效和仁慈的反应，而是行业中的一部分厂商利用政府权力为自己谋取利益的一种努力。

经济管制的这种解释，有时被称为"俘获"(capture)理论，施蒂格勒使用以工商企业或消费者为需求方，政府为供给方的供求分析经济框架。在这个理论中，一般说来管制被看成是由行业自己争取来的，因此，管制的设计和操作主要都是出于这个行业的利益。而且，对于私人活动来说，政府权力既是潜在的资源，又是潜在的威胁，如果购买了这些权力，就能让它作为资源发挥作用。

政府能够提供的利益(它产生对管制的需求)有：(1)直接补贴；(2)控制潜在竞争对手的进入；(3)控制替代商品和互

补商品的生产；(4)固定价格。在供给一方，他指出，使用政治过程的成本随着集团规模的扩大而增加，因此紧密的集团最有可能成为对管制的成功的“出价人”。因此，那些有能力购买管制的人就被允许使用政府的某些权力去获取利润，即重新分配财富以供他们自己使用。事实上，施蒂格勒的看法是，政治过程是以一种理性的方式建立和使用的。也就是说，政治过程被个人和集团利用来实现他们自己的欲望。

后来，佩尔兹曼(Peltzman,1976年)推导出施蒂格勒论文中隐含的数学模型。佩尔兹曼模型假定，管制者追求选票多数的极大化。佩尔兹曼随后指出，撇开别的不说，管制市场的均衡要求财产转移受益人的边际(政治)收益应当等于对其他人征税的边际(政治)成本。

A.2.4 波斯纳对施蒂格勒理论的扩展

当波斯纳(Posner)把施蒂格勒理论看作是理解管制的一项重要贡献的时候，他并没有明确表示支持，因为他觉得它(写于 1974 年)还没有精致到足以形成可以证伪的基本假设。因此，我打算把他提出的关于管制的经济理论——施蒂格勒提出的——和公共利益理论的一些论点作一个归纳。

和施蒂格勒一样，波斯纳也认为公共利益理论有某些不足之处，因为它不能解释为什么社会并不希望的后果会受到经济管制的支持。他举出的例子包括，铁路对第一个州际商法(Interstate Commerce Act)的支持，因为这项法规帮助铁路维持它们的卡特尔；美国电话电报公司(AT&T)希望管制能作为减少竞争的手段，以及货运公司和航空公司对它们本行业

实行经济管制的行为。

波斯纳还认为，虽然管制的类型符合公共利益理论，但是政府的管制机构却缺乏效率，从而导致管制的目标无法实现，这样一个看法也是有缺陷的。波斯纳反对这个看法，因为他发觉并没有足够的文献支持它。他认为，还没有一个具有说服力的理论能够解释为什么政府机构就应该比其他组织缺乏效率。他的推理涉及到对机构的领导和其他雇员行为的假设。他坚持认为，机构领导会对政府的立法和行政部门负责，而立法和行政部门则各司其职，并向管制机构提供经费。

政府机构内的雇员与追求在组织内能获得晋升，或在组织外增加他们销路的任何其他雇员一样，都是工作勤奋和诚实的。因没有找到垄断企业效率低于非垄断企业的令人信服的证据，所以波斯纳反对那种认为政府机构由于处于垄断地位而缺少效率的观点。我希望在本书的第6至8章证明，在这一点上他是错误的。总之，波斯纳只是假设X低效率等于零。

再回到他对施蒂格勒的评价上，波斯纳认为施蒂格勒的理论对管制提出了两个重要的见解：第一，可以把经济管制看作是根据供求法则配置的商品；第二，卡特尔理论能帮助我们把握(确定)供给曲线和需求曲线(的位置)。把管制看作一种商品，使我们的注意力转到管制的价值上，在其他条件不变的情况下，市场将把商品配置给那些愿意支付最高价格的买主。把管制看作商品，也使我们的注意力转到获得管制的成本上。

卡特尔理论告诉我们，一个卡特尔的价值越大，该行业产品需求的价格弹性就越小，进入这个行业的成本就越大而且时间越长。组成卡特尔的主要成本是达成协议和付诸实施

的成本。此外，卡特尔会碰到免费搭车的问题。卡特尔只有在销售者人数较少，而且他们的利益相对同质的时候，才容易达成和维持协议(把价格保持在竞争水平以上)。

和组成卡特尔一样，经济管制也有把价格提到竞争水平以上的效应。因此，管制的收益清楚地包含着这样一个重要的结果。获得管制的成本也是重要的，因为追求管制的行业必须同意采取这种形式，并避免出现免费搭车者(让不遵守管制其他方面规定的厂商不能索求管制的垄断价格)。厂商数目较少并且利益相对一致的行业，既适宜于实施管制也适合于组成卡特尔。

但是，行业管制和组成卡特尔之间也有不同。首先，在卡特尔不可行或者成本太大时，例如在那些集中程度较低的行业，对管制的需求就相对较大。换言之，当管制缺乏替代办法的时候，管制的需求就变得更大。其次，管制与政治过程的联系比较密切，而卡特尔通常却是厂商之间合作的产物。因此，能获得垄断控制的厂商至少有两种选择：或者组织卡特尔，或者要求管制。总之，管制的经济理论将管制描述成在以个人或集团充满理性和私利的人类行为对环境作出反应为特征的过程中，“提供给有效政治利益集团”的一种商品(服务)。当然，在恰当的供求条件下，管制的均衡价格和数量是会找到的。

A.2.5 寻租

我们说过，在均衡的市场上不存在X低效率。尤其是，如果受管制的市场处于均衡状态，那么X低效率就不可能是

那种管制的副产品。下一步，我们将考察一家厂商。标准教科书上的方法是假定厂商实现了成本极小化。X效率理论认为，偏离成本极小化就是存在X低效率的证据。另一方面，寻租理论把明显偏离成本极小化的事实解释成是人的理性行为的产物，从而完全排除了X低效率。同在市场均衡情况下一样，关于寻租行为的观点也假定不存在X低效率。

寻租观念依赖于租金的概念。租金是任何资源的报酬超出其机会成本的那部分。因此，租金就类似于经济利润。①在传统的经济理论中，在有秩序的——竞争性的——市场结构内追求经济利润（使收入流量的现值达到极大），既是个人行为的动机，又是促进公共利益的有用工具。这就是说，在有秩序的市场结构中追逐经济利润会产生社会剩余——即价值——因为它使那些经济利润降至零（在长期均衡中）。这种利润消失的原因是：它的存在（例如，由新产品的开发而产生）鼓励其他人从事寻求利润的活动。经济利润消失的过程（在理论上）很简单：流动与进入的自由使进入这个行业的企业的供给（资源）增加。因而，这个行业的产量上升，价格下跌，直到所有经济利润都消失为止。因此，利润的存在有助于吸引（配置）资源进入这个行业。因此，在有秩序的市场结构中寻求利润就导致配置效率。

但是，当资源是通过政治过程而不是市场过程进行配置时，结果就不同了。假定参加者的动机在政治过程中和在市场中寻求利润的过程完全一样，即都追逐个人收益。但产生的结果却是社会浪费而不是社会剩余。这个结论是从这样一个

① 寻求利润与寻租的异同，参见布坎南(Buchanan)1980年的著作。

假定推得的，即政治过程赋予某个个人或某个集团垄断地位，从而让其获取垄断利润。获得这种垄断地位的人和觊觎这种垄断地位的人都知道，政治决策并不是一成不变的。因此，这两个集团都有动机去配置资源以试图保持或得到政府的这种优惠。因而，政治过程会产生寻租行为。以这种方式配置掉的资源并不生产出具有市场价值的商品或劳务，相反，这种资源被用来造成有利于其花费者的财富分配。所以，虽然从个人观点看，寻租行为是合理的，但它却造成了社会浪费。因为寻租被假定是理性行为，所以由此必定可以推出，寻租以及特别是厂商因寻租而招致的较高成本，并不意味着厂商方面的任何低效率。因此，一旦寻租行为被假定是(例如)受管制厂商较高成本的原因，那末，X低效率是政府管制的结果这个命题就被假设掉了。

A.2.5.1 塔洛克论垄断的福利成本

塔洛克(Tullock)1967年论述垄断的福利成本的论文是一篇令人感兴趣的文章，这不仅因为它本身的缘故，而且因为它表明了管制的这种“正统”观点与X效率观点之间的差异。塔洛克一开头就引用罗伯特·芒德尔(Robert Mundell, 1962年)关于对垄断力量引起的福利损失(市场配置低效率)的过小估计会让一些人得出经济学并不重要这种结论的一段痛惜话。芒德尔对此作出的回答是重新考查“赖以进行这些估计的工具”。塔洛克然后引用了莱宾斯坦1966年那篇最早论述X低效率的文章，在这篇文章中，莱宾斯坦讨论了为什么小的配置损失(由福利三角形度量)可能并不是唯一的，或并不一定

是最大的由垄断力量造成的福利损失。塔洛克对莱宾斯坦意见的回答则是，否定其他那些损失属于非配置损失的可能性，而主张沿着芒德尔提出的思路，对所使用的工具进行重新估价，并只对配置(低)效率保持关注。这换句话说就是，塔洛克假定X(低)效率并不存在。从一开始就明确这一点看来是有用的。

塔洛克和其他人对寻租所作的研究，我们可以用图2-7来说明。图2-7可以说明政府福利成本的几个例子。第一个是货物税的例子。我们可以想象，倘若没有这种税，不变成本保持在C_0水平上，国内价格为P_C，结果将导致产量Q_C。可是，征收货物税把价格提高到P_m，产量减少到Q_m。福利成本为面积ABC，而面积P_mABP_c则从这种商品的购买者转移到公共基金的接受者。但是，塔洛克假设提高价格得到的收入被例如挖掘并不通向任何地方的隧道完全浪费掉。在这种情况下，

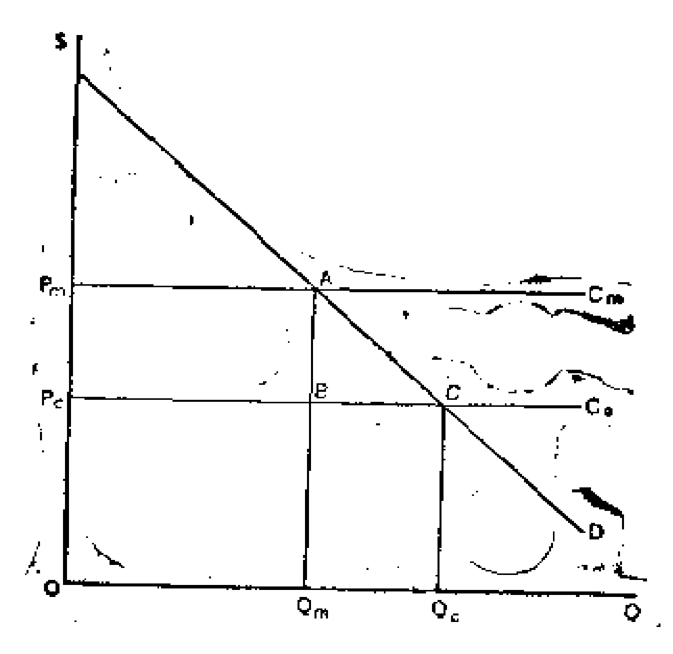


图2-7 税收、关税和寻租的福利成本

政府税收计划的全部福利成本就是整个P_mACP_c面积。也就是说，在这种情况下，不但出现市场配置低效率ABC，而且还有并不产生任何收入的支出P_mABP_c。这里不涉及任何转移，而只有稀缺资源的浪费。

第二个是关税的例子。我们假定成本保持不变。没有关税时在P_c价格上可以购买的单位是Q_c。征收关税后，在P_m价格上的购买单位为Q_m。关税的福利成本是三角形ABC，而由关税造成的较高价格被认为是从一个集团向另一个集团的转移(P_mABP_c)。

但是，如果我们假设，国内生产者的效率低于外国生产者，情况又会怎么样呢？如果商品的国内成本为C_m，而商品的国外成本为C_c，那末要是没有关税，低成本的外国生产者将按价格P_c为整个市场生产Q_c。征收禁止性关税使需求降低到Q_m，并使效率较低的国内生产者得以控制整个市场。在这种情况下，不但存在市场配置低效率ABC，而且政府的关税政策使生产向处于比较劣势国家的国内生产者移动。这种生产向低效率——高成本——生产者的移动(shift)，并不产生转移(transfer)而是产生由面积P_mABP_c度量的浪费。再次地，在这种情况下，政府关税的福利成本是整个面积P_mACP_c。在货物税和关税这两个例子里，政府政策的整个福利成本都超过了由垄断力量引起的福利损失的传统度量——面积ABC。

在他论文中部分论及犯罪的那一节里，塔洛克继续这样说道，政府征收关税常常是因为某个利益集团为使政府征收这种关税而付出的支出(假如是合法的)所造成的压力。据塔洛克所说，与一般的极大化原则相一致，这些支出将一直持

续到它的边际报酬等于它的机会成本为止。但是，在同时会有许多集团都在试图影响政府，因此，这些支出中有许多仅仅起到了抵消其他人支出的效果。这些(寻租)支出并不创造财富，因而是整个垄断(或者获取垄断地位的努力)造成福利损失的一部分。这里，我们假设，这种寻租过程将消耗所有的租金或利润。

虽然这种寻租支出难以数量化，但塔洛克相信，它们可能相当可观，因为这种支出的潜在报酬很大。这就是说，寻租支出是从政府优惠中所获得的预期或潜在报酬规模的函数。必须强调，这些支出虽然从社会的观点看是一种浪费，但从支付它们的利益集团的观点看，它却是“理性的投资”。还应强调，如果我们假设这些寻租支出是理性的——因为是边际相等的决策——那末X低效率就已经被假设掉了。因此，根据定义，政府管制就不可能造成X低效率。它实际所做的只是为雇员进行在社会看来是浪费的私人理性投资予以鼓励。必须强调指出，这里的厂商不是(X)低效率的。在产生市场配置低效率和其他形式的社会浪费(寻租支出)方面，市场被取而代之了。但厂商迄今仍被假设是成本极小化者，它还是有效率的。最后，读者应该认识到，虽然X效率理论与生产成本的关系更为紧密，但是塔洛克选择强调的却是寻租支出。

根据图2-7，并且假定全部租金都被耗尽，塔洛克的主张可以解释为这样一种假设，较高的成本——面积 P_mABP_c——只是由于寻租支出所致。事实上，对寻租理论作出重要贡献的三篇早期著述——克鲁格(Krueger，1974年)；波斯纳，1975年；塔洛克，1980年——都认为租金恰好耗尽。

A.2.5.2 克雷恩和札德库希

许多作者，其中包括克雷恩和札德库希（Crain and Zardkoohi，1980年）都对这个问题进行过经验研究。这些作者的看法是，生产中存在的X低效率，通过减少可获得的寻租投资基金，包含着一种较低利润的机会成本。因此，他们假设在一个"寻租社会"里，厂商有（强烈的）动机追求X效率，因为只有这样，它们的寻租投资基金的供给才能足以使它们有能力影响政府以满足它们的利润需要。因此，X低效率不是一种免费商品。而且，较多的X低效率只是意味着较少的另一种类型的寻租低效率。因此，社会面对的是两种类型低效率之间的权衡。

另一方面，具有较多X效率的厂商会产生参与其他类型的寻租低效率的基金。这种情况如图 2-8 所示。图 2-8 把面积P_mABP_c分作三类：(1)面积 1 ——$P_mAB'P_m'$——代表利润；(2)面积 2 ——$P_m'B'B''P_m''$——代表寻租支出；以及(3)面积 3 ——$P_m''B''BP_c$——代表纯粹的X低效率。此外，读者还可以看到例如在寻租支出的成本超过它们对厂商（或集团）的价值的意义上，面积 2 可能包含X低效率。

在他们对X低效率和寻租（有关他们的经验研究结果的细节，可在他们的论文集中找到）所作的分析中，他们的结论是：(1)纯粹X低效率（面积 3 ）能够存在；(2) X 低效率的变化假定被寻租支出的等量变化所抵消（面积2）；(3)因此，X低效率即使确实存在，它也不会导致全部福利的净变化。换言之，即使它存在，也被假定没有任何重要性。

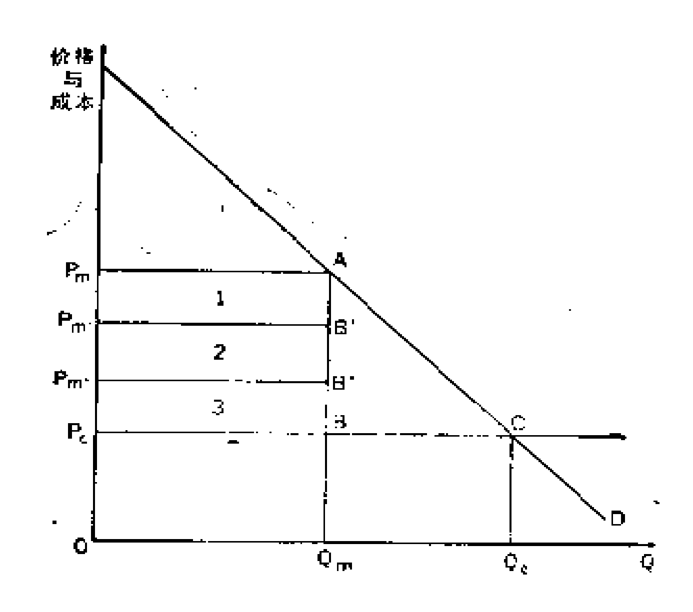


图 2-8 把与垄断力量有关的较高成本分成为：
利润、寻租和X低效率

A.2.5.3 波斯纳(1975年)论寻租是一种竞争活动

在获得垄断地位是一种竞争活动和理性行为的假定下，波斯纳正确地接着假设，在边际情况下，获得垄断的成本等于成为垄断者的预期利润。这是标准的极大化假设。就是说，所有利润全都耗尽。

他把航空业中的非价格竞争作为例子来讨论，如图 2-8 所示。波斯纳说，如果民用航空委员会在航运价格下面规定的最低价格(P_m)超过提供航空运输的边际成本(C_o)，那末这个行业所赚的利润就等于面积P_mABP_c。可是，如果利润是可获得的，非价格竞争假定并不受到管制，那么厂商将在非价格竞争上扩大资源使用量，直到边际成本提高到C_m，以及经济利润消失为止。

此外，如果非价格竞争也受管制和被禁止，那末航空公

司的预期利润就会较大，因为利润不会因为非价格竞争而消失。由此"必然出现"这样一种情况，厂商将花费更多数量的货币以图拿到经营航空公司的执照。其结果就是，在这种情况下，经济利润同样会全部耗尽。最后，读者应该知道，在波斯纳看来较高成本不该被假定是X低效率的证据，而应该被假定从厂商观点看来是理性的投资。(回忆一下，上文对波斯纳观点的讨论表明，他是把X低效率的存在假设掉的。)

A.2.6 结论

用正统微观经济理论来评价X(低)效率概念是困难的，因为这种理论的假设（包括极大化）本身就不允许那个概念具有真正的意义和(或)重要性。把X(低)效率作为政府管制的后果来估价，如果我们用的是现在流行的那几种运用正统微观经济学范式的管制理论的话，那么肯定是不会有什么结果的。估价管制的潜在X低效率效应，要求我们使用至少与X效率理论相容的理论框架。我们将在第4章和第5章再回到这个问题上来。但是，首先，我们要对莱宾斯坦关于X效率理论的最早论述，以及在这一段时期内发展起来的其他(非利润极大化)厂商理论的概况作一番叙述。

3 X效率：学术背景和理论介绍

3.1 导言

在上一章，我们评论了关于生产、成本和垄断力量福利成本的微观经济理论。我们知道，传统的微观经济学假设企业是利润极大化者和成本极小化者，效率指的是市场配置效率。这就是说，企业从给定投入获得最大的产出，为给定的产出耗费最小的成本。此外，采取这种分析框架导致若干当代的管制理论假定不存在（由这种管制造成的）X(低）效率。现在是对这些假设提出疑问的时候了。我们特别要对企业试图实现利润极大化和成本极小化这个假设提出疑问。我们还要说明，哈佛大学教授哈维·莱宾斯坦在向这些假设提出疑问的时候，建立起了一个假设存在非配置型效率的理论。1966年，他把这种效率称为X效率。正是来自工业化国家和欠发达国家的数据资料，引起他对这些假设产生了疑问。这些资料可以分为两种类型：第一，说明企业不是内部有效率的资料；第二，说明企业没有实现利润极大化的资料。

当时，向微观经济学的某些假设提出疑问的，并不只是

莱宾斯坦一个人。例如，我们前面所论及的塔洛克论垄断力量的社会成本的著述，就是在莱宾斯坦论文问世后次年出版的。本章将考察这样一批观点相同的作者的意见，他们都假设企业试图使利润以外的某样东西实现极大化。这就是说，组织的目标函数是“复合的”。然后，我们将阐述促使莱宾斯坦最初提出X效率理论的资料。这样，我们不但可以初步懂得X效率理论，而且还可以懂得它与其他厂商理论有何不同。

3.2 复合的目标函数

3.2.1 导言

可以肯定，在经济学中，关于单个厂商和厂商集团具有非利润目标的看法，并不是什么令人吃惊和感到新奇的观念。艾尔弗雷德·马歇尔（Alfred Marshall）在他的《工业与商业》(1932年)一书里，伯利(Berle)和米恩斯(Means)在他们经常被人引用的《现代公司和私有产权》(1932年)一书里，凯恩斯(Keynes)在他的《劝说集》(1935年)里，都谈到过现代厂商决策者的非利润目标。希克斯(Hicks，1935年）曾争辩说，“宁静的生活”也许是垄断利润的最好形式。当代作者讨论这个问题的更多，包括有科尔（Cole, 1959年)、戈登（Gordon, 1961年)和巴纳德(Barnard,1962年)。这些作者认为，除利润外，权力、威望、胜任职业、薪金、安全和地位中的一个或几个都适合成为经理人员的主要目标。

在这一章里，我将阐述其他六位经济学家对明显是非利

润极大化的厂商行为所作的反应。论述的次序是：蒂博尔·西托夫斯基（Tibor Scitovsky，1943年），威廉·鲍莫尔（William Baumol,1959年），罗宾·马里斯（Robbin Marris，1963年、1964年），奥利弗·威廉森(Oliver Williamson,1964年)，小约瑟夫·蒙森和安东尼·唐斯(R. Joseph Monsen,Jr. and Anthony Downs，1965年)。

3.2.2 西托夫斯基

西托夫斯基的论文，虽然是在莱宾斯坦首次论述X效率20多年之前发表的，但他的研究方法却是独创性的，并成为其他人的范式。在西托夫斯基的模型中经理的效用函数包括收入与闲暇这两个变量。虽然这里的论述与西托夫斯基论文中的阐述不尽相同，但是下面的叙述确实体现了他模型的大部分精神实质。在图3－1中，纵坐标衡量利润，横坐标衡量不活动性(inactivity)。利润在不活动水平OA上达到极大。为了使企业家同时实现他（或她）自己的效用和利润的极大化，要求他(或她)的无差异曲线在点B上与这条利润线相切。这是根据效用极大化规则——效用在无差异曲线斜率与预算线(在这里是利润线)斜率相切的地方达到极大——推得的。但是，为了使效用在点B上实现极大化，企业家必须有一条在利润水平OC上完全水平的利润与闲暇的无差异曲线(CD)。也就是说，闲暇对利润的边际替代率必须等于零；换言之，不管企业家会获得多少闲暇，倘若利润低于OC,他就不能再维持在无差异曲线I_1上。看图3－1的另一种方法是，在利润水平OC上，企业家努力的供给是无限的，而他闲暇的边际效用

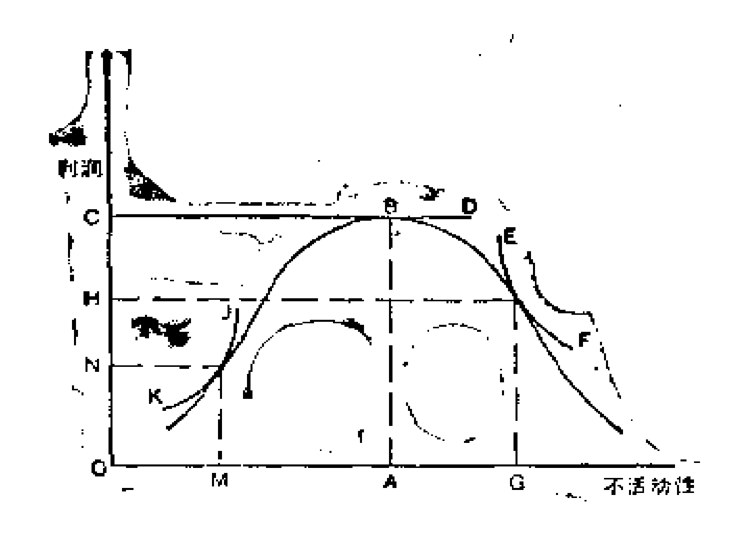


图 3-1　效用和利润极大化

却等于零。

与无差异曲线 CD 相比，让我们来考察一下无差异曲线 EF。这条曲线表示企业家愿意以闲暇替代利润。这样，利润和闲暇的边际效用都大于零。在这个例子中，企业家将在闲暇水平OG和利润水平 OH 上实现他(或她)的效用极大化。另一方面，如果利润或闲暇都是低档商品，以致像无差异曲线 JK那样向右上方倾斜的话，那末企业家将在较低的利润水平 OM 和闲暇水平 ON 上实现他（或她）的效用极大化。当然，在这两种情况下(无差异曲线 EF 或 JK)，努力的供给都不是无限的。

虽然企业家被当作是其工作努力以无差异曲线 CD 为特征的那种人的代表，西托夫斯基仍告诫说，这种工作心理“不一定适用于每一个企业家，甚至对于有代表性的企业家来说也未必正确。”虽然它没有描述利润极大化，但是，按照西托夫斯基的说法，它却“使经济分析变得十分简单”。

3.2.3 鲍莫尔

威廉·鲍莫尔的模型是他与一家商业咨询公司合作的结果。他假设，"典型的"寡头垄断者的目标，是在厂商获取某个"最低"水平利润的约束下，实现销售量（总收益）的极大化。鲍莫尔并不认为这个目标与理性行为有什么不一致。他说，"人们的目标是什么就是什么。非理性行为无疑应该这样定义，即它应存在于那种使自主目的更难达到的决策模式，而不在于选择出于某种理由被认为是错误的目的。"实现销售量极大的愿望，部分是因为销售量与经理的薪金存在正相关关系。其次，之所以希望实现销售量极大，是因为销售量下降会带来许多不利，其中包括丧失顾客(随波效应)、丧失在资本市场的融资能力、丧失分销商和更多繁复的行业关系。鲍莫尔报告说，当利润和销售量彼此发生冲突时，企业家几乎总是宁愿保持或增加销售量。虽然无利可图的销售最后会终止，但这种事情仅在"再三深思和踌躇"后才会出现。

销售量极大化要求关心未来的销售，并且关心为增加未来销售所必需的扩展筹措资金。因此鲍莫尔把最低利润约束定义为支付这种扩展费用所必需的那个利润水平。(这种扩展的费用并不全是用留存收益支付的。因此，利润还必须高到足以能从资本市场上吸引到足够的资金。）因此，"中等利润水平"可能是达到长期销售极大化的"最优"水平。一旦达到这种最低利润水平，销售极大化就成了厂商的主要目标。鲍莫尔报告说，在他的经历中只碰到一次这样的情况：新闻界引用了一家大公司董事长的话说，他要使他的公司成为这个

行业中最能赚钱而不是最大的公司。"饶有兴趣的是，这位董事长过去曾是一位研究学术的经济学家。"对于长期销售量极大化是不是与长期利润相一致这个问题，鲍莫尔补充说，在存在诸多有不同权威的权力中心的大公司中，他们并不经常讨论和明确制订目标，因此，长期目标概念就变得很难理解。

3.2.4 马里斯

对于罗宾·马里斯来说，现代公司——寡头垄断——的特征是所有权与财务分离。在马里斯的模型中，经理们追求效用的极大化，其效用则是厂商增长和他们自身安全的一个函数。增长可用资产的增加来衡量，并反映了从与不断发展的大厂商的联系中得到的满足、薪金、权力和威望。安全则以厂商的估价比率来计量，也就是厂商的市场价值与其资产帐面价值的比率。随着估价比率下降，厂商越发可能成为被接管的目标，这对经理层的地位是一种威胁。马里斯认为，就这种效用函数来说，经理层将在安全约束下追求增长的极大化。也就是说，经理层的首要目标是使被接管的机会降到最小，做到这点以后，他们才追求厂商净资产增长的极大化。(这些厂商被称为"经理型厂商"，而那些由忽视增长的经理们管理的厂商被叫作"古典型厂商"。古典型厂商的首要目标是股东财富的极大化。)

可是，较多的安全与较高的增长并不一定相容。首先，由减少资本报酬率所取得的较高增长"降低了"利润。增长不一定与利润相容的另一个原因在于，增长不一定与效率相容。

同假设所有工人和管理人员都可以由另一批工人和管理人员完全替代的较为传统的新古典模型相比较，马里斯假设组织的新成员，不管多么胜任，在他们熟悉该组织的文化以前是不可能完全有效地进行非例行决策的。因此他假设，经理层的平均决策效率是下面两种要素的函数：（1）与“老”成员平均效率相比的新成员的平均效率；（2）新成员总数在组织的全体员工数中所占的百分比。由于厂商的增长增加了新成员在全部成员中所占的比重，因而增长将降低平均任职期乃至平均效率。虽然厂商增长最初可能通过引入新观念使效率增加，但最后还是会经过上述机制而降低效率。

马里斯似乎是假设经理层会受被接管威胁约束的首批经济学家之一。然而，虽然新古典理论假定非利润极大化总会导致接管——假定的前提是完全的资本市场——但是，马里斯却假定资本市场是不完全的，赋予经理层在经营中不受惩罚的某种自由处置权。其解释部分取决于缺乏有能力的“侵入者”。新古典理论认为工人与管理人员都是同质的假设会导致这样的理论，即在任何一个既定规模、既定的劳动力、技术和商业信誉的企业里工作的任何经理人员，就应该能提供相同的利润。这就是一种完全知道生产函数的假设。撇开别的不谈，它假定，对厂商的熟悉和经验不是生产率的重要决定因素。马里斯反对这种假设，认为这些因素至关重要——它们影响生产率和利润。

潜在的公司侵入者必然进入一家具有既定的劳动力、管理人员和商业信誉的企业。每个资本家都能够提供足够“高明”的经理去赚取比先前的经理更多的利润吗？他们能够使用先前业主留下来的劳动力和商业信誉做到这一点吗？马里

斯假定大多数资本家都不能提供必需的经理人员，这就限制了来自资产者阶级的侵入者人数。管理专家虽然有这种能力，但是通常都缺乏资金。而且，其他管理专家也可能是以增长为导向的。因此，容许的自由处置行为既可由经理也可由市场加以确定。马里斯无视经常存在的被接管威胁，而是假定，最有效率的厂商增长最快，效率较低的厂商增长较慢，最没有效率的厂商将被接管。

3.2.5 威廉森

奥利弗·威廉森把非货币价值并入他的带有"支出偏好函数"的动机与行为模型之中。威廉森假定，经理们管理企业，目的是使他们自己的效用达到极大，而其效用则是职工、报酬和可自由处置利润的函数。经理们试图在税后利润大于或等于某个最低水平的约束下，实现这种效用的极大化。在威廉森写论文的当时，人们的习惯看法是假定经理们对各种类型的成本或支出没有差异：1美元就是1美元。在他的支出偏好模型里，经理们对成本的态度是不对称的，这就是说，对某些类型支出的偏好胜过对其他类型支出的偏好。归入得到偏好的支出范畴，是那些对经理产生效用，但不能对生产率有积极影响的支出。

用于职工的支出——一般管理和销售支出——会在几个方面为经理们提供效用。第一，在一个固定规模的企业里，职工的增加是获取晋升的一条途径。第二，由于职工人数增加是晋升的一条途径，因而它也是提高薪金和更多地支配别人的一条途径。第三，职工增多常常会减少经理对企业生存的

忧虑，从而使他们的安全感增加。报酬代表的是经理人员可以自由处置的薪金和津贴，也就是说，报酬代表"租金"。因此，报酬不会提高经理层的生产率。相反，它们代表的是对由经理层获取的垄断报酬的分配。像铺地毯的大办公室、费用帐户、特别秘书服务或交际预算诸如此类的报酬，会通过满足对这些东西的欲望以及有助于提高声望和地位而使效用增加。有些报酬是以薪金形式获取的。可是，为了避税，它们被分作薪金和津贴。可自由处置的利润据说是由税后利润减去股东要求的最小利润的差额来衡量。可自由处置的利润是报酬和成就感的来源，所以它会产生效用。

利润与支出偏好——在这种情况下，是用于职工的支出——之间的权衡如图 3-2 所示。可自由处置的最大税后利润出现在需要职工人数OB的点A上。但是，给定无差异曲线I_1，效用在职工数为OD，可自由处置的税后利润为 CD 的点C上达到极大。这就是说，效用极大化需要较多的职工，因而需要比达到利润极大化所要求的更多产出。因此，经理们被认为会转移资源，在这个例子里，是将资源移用于职工，而不是把它们报作利润。虽然给定 I_1 后，利润极大化和效用极大化不能兼容，但是，必须清楚，倘若利润和职工间的无差异曲线如EA所示，则效用也会在点A上实现极大化。也就是说，利润对职工的边际替代率为零，隐含着职工的边际效用等于零。这种情况意味着，职工之所以具有价值，只是因为它对利润有贡献，或者因为它对经理层不产生任何效用。由于在威廉森的"支出偏好"模型里，前者已被排除，而后者不可能出现，所以把无差异曲线画作I_1 会使利润极大化和效用极大化不相容。

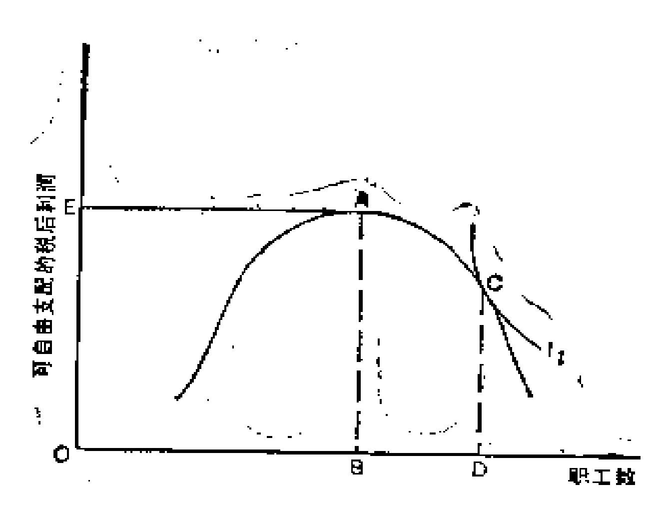


图 3-2　自由处置行为、利润和支出偏好

威廉森认识到，现代企业组织是一个联合体，它是一个组织，而不是像比较传统的教科书所描述的那样仅仅是一个"概念"或"代表性厂商"。但是，他补充说，只是在职位和收入受到威胁的危机时刻，它才是平等个人的联合体。在非危机时期，经理层成为这种联合体的主要成员。这使经理人员能占有其他人得不到的信息，从而使经理层获得决策权。虽然被接管的威胁会影响经理层的行为，但它只是一种不完全的制约。因此，投资者的潜在权力"大部分未经行使就丧失了"，只剩下资本市场成为投资者赖以控制经理的不完全机制；因而，结论就是经理们的动机和行为十分重要。

3.2.6　蒙森和唐斯

蒙森和唐斯一开始就指出，传统的价格理论只研究一种类型的厂商——业主管理的小型厂商。因此，这个理论把厂商看作是单个的人，具有已知的和有序的偏好，并有能力去实现这些偏好。与此相比，蒙森和唐斯表明，公司的规模和所

有权形式都会影响行为，特别是那些其主要目标并非利润极大化的大型管理型企业。（它们至少有1000名雇员，其经理层在企业里没有具支配力的股权利益。）蒙森和唐斯假设，在这些企业中所有者想要的是稳定的股息收入和逐步上升的股票价格，而经理层试图实现的却是使他们的终身收入极大化。

如同在马里斯与威廉森的模型里那样，蒙森和唐斯也假定资本市场是不完全的，因而会降低所有者控制经理追求利润极大化的能力。这是因为，相对来说，所有者对于企业所面临的各种可供选择的方案，事前是不太了解的。他们只在事后才知道厂商的绩效，因而实际上并不能断定利润是否达到了极大。此外，所有者也无法"精确判断经营质量上的微小差别"。由于资本利得税加重了他们的"负担"，所以在边际上的股票转换减少了，"不同股票之间的竞争力量，在理论上可以预期会向最高经理层施加压力，使他们让企业的股票价格上升率达到极大，但实际上却会因无知和税收结构而被严重削弱。"虽然潜在股东不购买股票对经理们起了一种"潜在制约"的作用，公司的侵入者也增加了这种制约力量，但是这些制约作用还是不完全的。因此，只要股票的价格符合"令人满意的增长"标准，股息也不减少，经理们将有理由觉得他们的职位安全无虞。

对于经理们来说，他们的收入包括货币报酬和诸如闲暇、声望与权力的非货币报酬。假定经理们把后者看得与前者同等重要，并追求使他们的终身收入和美元价值实现极大化，至于他们在大型管理型企业里的实绩，这些组织的官僚性质一方面使得他们可以掩盖信息，以至只有对他们有利的信息才被披露；另一方面，又使得他们可以只"积极"落实他们的

那部分要求。希望使他们终身收入实现极大化的经理们不但会掩盖信息，而且会指挥企业，使它支付稳定增长的股息，并让它的股票价格稳步上升。此外，经理们还是风险回避者，他们竭力避免公司收益的多变，因而会减缓公司的增长。换言之，大型企业尤其是大型管理型企业，它们的效率将低于业主管理的小型企业。这就是说，前者更可能赚取不到最大利润。与此同时，经理们会利用他们的职位，谋取自己个人的声望，并获取诸如费用帐户这类不像薪金那样看得见的额外津贴。与威廉森的模型类似，蒙森和唐斯的模型也表明，效用(终身收入)极大化和利润极大化不能相容。

3.3 X效率理论：简介

3.3.1 导言

上面讨论的那些厂商理论假定，厂商都除了追求利润以外还追求某种东西的极大化。这里关键的一个词是极大化，如果厂商是追求除利润之外的极大化的话，这些著述者就描述了厂商的这种行为。这些著述者所作的这些贡献，丰富了我们关于厂商的知识，并使微观经济学变得更加切合实际。(可以相信，后一点是这些著述者的目标之一，而且他们在这方面取得了进展。)可是，这些理论仍然把厂商的要素视为"黑箱"。这就是说，他们先假设存在一个并非利润的目标函数，然后说明这个函数如何对厂商产生影响。或许，这是研究微观经济学的最佳方法。但是这里要强调指出的是，在某些方面X效率理论与这种方法不同。其中之一就是，X效率理论

试图更广泛地研究个人工作(努力)选择的心理基础，以及个人如何受周围集体影响的问题。第 4 章和第 5 章将集中讨论这些问题。在本章中，我将阐述促使X 效率理论发展起来的一些资料，和莱宾斯坦1966年论文的概要。促使X 效率理论发展起来的资料有两种形式：能明显说明厂商并非内部有效率的资料和说明厂商因为没有遵循边际分析的行为规则，所以不会使它们的利润实现极大化的资料。

3.3.2　内部效率

内部效率：微观资料

1962年，彼得·基尔比(Peter Kilby)提出国际劳工组织的生产率论证机构对几个国家不同行业调查的报告结果，被调查的这些国家有缅甸、希腊、印度、印尼、以色列、马来亚、巴基斯坦、新加坡和泰国。这次调查结果表明，“简单地改变一下一家工厂生产过程的实际组织”就能够使劳动生产率有相当大的提高，单位劳动和资本的成本有相当大的下降。由于劳动生产率用Q/L度量，单位劳动成本用LW/Q度量，所以对于既定的工资率而言，生产率上升就意味着单位劳动成本下降。这些“简单的改变”包括工厂布局的改变，以及机器的使用、原料的管理、整个生产的流程、浪费的控制、职工工资支付的方法、工人的训练和监督的改变。上述所有 9 个国家的情况表明，劳动生产率平均提高75％，单位劳动和单位资本的成本就平均下降约35％。

基尔比还报道了尼日利亚 5 家橡胶厂的资料。在这个例子中，所有这些工厂都位于同一个地区，它们都在同一市场上

销售产品，在同一市场上雇用劳动力，所有工厂都使用具有相同实物布局的相同资本。尽管如此，它们的单位劳动需要量(总人时/总产量)和单位资本需要量还是分别相差134%和270%之多，这些厂商中最有效率的，即单位投入需要量最低的是那些由希腊人拥有和经营的厂商；而效率最低的是尼日利亚的私人所有企业。很清楚，基尔比的资料告诉我们，产量与成本不光由技术决定，相反，厂商组织应该被视为产量和成本的重要决定因素。

L·罗斯塔斯(L.Rostas) 1962年对美国和英国 31 家制造业工厂的相对劳动生产率所作研究的结果表明，美国劳动生产率平均比英国劳动生产率高120%。这个差距并不一定令人吃惊。但是，他特别提到，“在美国和英国许多设备相同的行业(或企业)里，英国和美国工人的人均产量却仍有很大的差距。”其他的因素还包括工厂布局；生产流程；照明、供暖和其他的“工厂内部条件”；工作小时数（他报道说在美英两国当时的工作小时水平上，工作小时与生产率之间呈相反的关系）；工资支付方法（计时或计件）；譬如得自时间和动作实验的工作方法简化，工人流动率和促动因素。他说，“更多地了解这些‘其他’因素，知道它们在多大程度上影响每个工人的产量十分重要。这些因素独立于生产机制或生产技术，它们部分属于‘组织’因素，部分属于影响工人的意愿和能力的因素。”

罗斯塔斯没有提供有关这些“其他”因素的任何数量估计。可是，C·F·普拉顿(C.F.Pratten, 1976年)在他对多国公司工厂间的生产率差异所作的研究中，的确提供了这方面的估计数字。虽然普拉顿是在莱宾斯坦最初提出X效率理论

之后10年才发表其研究成果的，但是，根据罗斯塔斯的研究，普拉顿的研究看来还很重要。普拉顿报道了经济和行为因素对多国公司在美国和德国的工厂、英国和法国的工厂、英国和北美的工厂之间生产率差异的影响的估计数字。按平均数计算，英国工厂中的劳动生产率要比德国工厂低27%，比法国低15%，比北美低50%。经济因素包括产出率、生产流水线长度、资本、产品组合和生产能力利用率。行为因素包括罢工和劳动准则(包含影响生产率的劳资关系的作用)。对于英国和德国工厂之间的差距，经济因素可以"解释"13个百分点，行为因素可以"解释"12个百分点。对于英国和法国之间的差距，经济因素和行为因素可以分别解释9个百分点和5.5个百分点。对于英国和北美之间的差距，经济因素和行为因素可以分别"解释"35个百分点和11个百分点。

弗雷德里克·哈比森(Frederick Harbison，1956年)在其经济发展过程的案例研究中，也着眼于"组织"的贡献。他认定，组织超过从事管理、计划、协调、执行和监督的诸多个人的总和。组织这个要素不但体现了这些功能，而且还是这些功能的"完整综合"。哈比森补充说："在技术不变的情况下，雇佣劳动的组织在劳动生产率的决定中也许是首要的因素——即是居支配地位的力量。"因此，组织这个要素是企业中"经常"观察得到的劳动生产率差异的重要决定因素。例如，哈比森报道说，埃及两家彼此相距不到1.5英里的炼油厂，其中一家的劳动生产率在几年里都是另一家的2倍。"但是近来那家效率差的炼油厂，因为在全新的经理层领导下，虽然还使用同样的劳动力，效率却已经开始有了令人瞩目的改善。"

可以这样解释，这些报告说明这些企业并没有按照它们的生产函数进行生产。在图3-3a中，企业是在点C而不是在点A或点B上进行生产。因此，可以说企业是按以下两种情况中的一种生产的，要末是在既定的投入(L_2)下，产量少于最大的产出(Q_2)，要末是使用超出技术所必需的最低投入(L_1)生产既定的产出Q_1。进一步说，这就意味着它们的实际

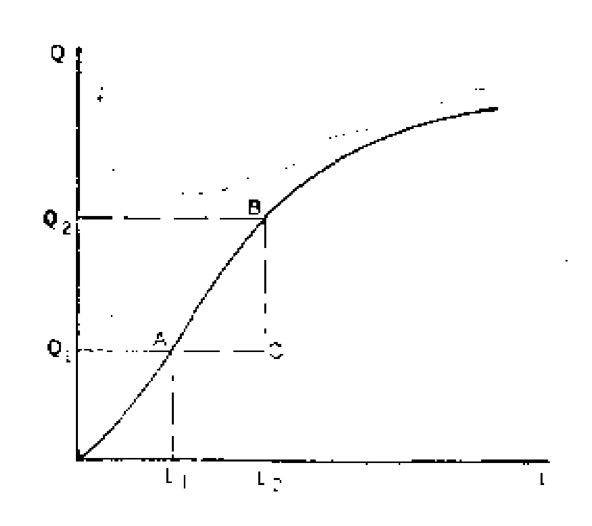


图 3-3a　X低效率和短期生产函数

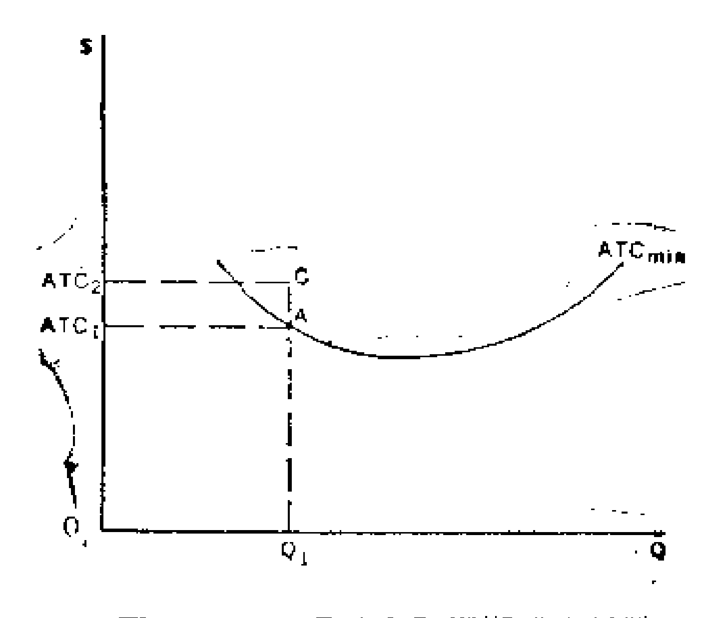


图 3-3b　X低效率和短期成本函数

成本高于它们技术上必要的最低水平。在图3-3b中，生产Q_1的企业应该在点A上进行生产，由此招致的成本是ATC_1。但是，因为它们实际使用的投入超过必要的投入，所以它们的实际成本也高于必要的成本(投入价格既定)。因此，该企业将按成本ATC_2在点C进行生产。

生产率和增长：宏观数据

另外几项对宏观经济行为的研究，可以看作是上述研究的补充。1957年罗伯特·索洛(Robert Solow)发表了他关于美国非农业私营经济部门总生产函数的研究成果。他研究了1909—1949年这样一段时期的资料。他的结论可以概括如下：(1)资本存量增加1%，产量就会增加0.35%；(2)劳动力增加1%，产量就会增加0.65%；(3)"技术变化"使产量平均每年增加1.5%。索洛把技术变化定义为生产函数的任何一种移动，即既定资本和劳动投入的生产能力变化。技术变化包括的要素有"工作放慢、工作加快、劳动力教育程度提高以及诸如此类的事情"。换言之，技术变化包括除传统的资本和劳动投入之外的所有影响生产的要素。这样看来，它应该包括上面所定义的"组织"要素。

1959年，索洛发表了另一份研究1919—1953年美国非农业私营经济部门的报告。在报告中他说，资本和劳动增加1%，产量就会分别增加0.30%和0.70%。但是技术变化却使产量平均每年增加2.5%。

奥德·奥克鲁斯特(Odd Aukrust, 1959年)发表了一份类似的研究1900—1955年间整个挪威经济的报告。他的资料说明：(1)资本增加1%，产量就会增加0.20%；(2)劳动力增

加1％，产量就会增加0.76％；（3）由于“组织”的改进，产量每年增加1.8％。奥克鲁斯特使用与索洛给技术变化下定义的相同方式给组织下了定义。他特别提到了经理人员与其他雇员的专门技能、工作动机、工作的社会风气和国际环境。

奥拉维·尼塔莫（Olavi Niitamo,1958年）发表了另一份研究1925—1952年芬兰制造业部门的同样报告。他的资料表明：（1）资本增加1％，产量就会增加0.26％；（2）劳动力增加1％，产量就会增加0.74％；（3）组织改进使产量每年增加1.2％。这些研究结果综合起来，证明存在着一个对生产有重大影响的非传统因素。这些发现并不是X低效率的证据。这里提到它们，是因为莱宾斯坦在提出X效率理论的时候，了解并考虑到了它们。

3.3.3 利润极大化

从20世纪30年代后期开始，许多研究报告都指出，厂商并没有实现利润极大化，或者厂商的行为方式并不能使它们实现利润极大化。在这些研究中，霍尔和希契（R.Hall and C. Hitch,1939年）的研究是首批之一。霍尔和希契是来自英国剑桥大学的一批经济学家的成员，这批经济学家把研究企业的实际行为，作为他们对商业周期研究的组成部分。霍尔和希契的研究以38位企业家的回答为根据，其中有33人是从事制造业的。

经济理论假设，企业产量要达到边际收入（MR）等于边际成本（MC）时为止，企业将根据需求决定它们的价格。与此相对照，在这38位企业家的回答中，最“令人惊诧之处”在

于，企业并不使MR与MC相等。此外，它们也不用根据需求来决定价格。实际上它们实行的是“完全成本定价”，也就是，它们力图估计出平均总成本，然后开出一个包含这些成本以及利润加成的价格。换句话说，“在定价时，他们试图应用的是我们称之为‘完全成本’的比较粗糙的方法，如果利润极大化完全是应用这种方法所导致的结果，那么利润极大化就只能说是意外的(或可能出现的)副产品。”大部分企业都声称执行完全成本定价原则：75%的垄断和寡头垄断企业，62%的垄断竞争企业，73%具有这两种不完全竞争形式特征的企业，都声称“严格”或“正规”执行完全成本定价方法。大部分企业还回答说，它们执行完全成本定价方法是因为它会导致“正确的价格”。霍尔和希契还报告说，企业家使用完全成本定价政策是因为他们并不知道产品的需求曲线和需求的价格弹性！没有这种信息，就不可能根据边际分析来决定价格。而且，企业家缺乏需求曲线方面的信息，就意味着他(或她)缺乏边际收益的信息。因此，霍尔和希契说：“他们使工资与雇佣劳动的边际产量(或边际收益)在短期有密切关系的假设也变得不再成立。”

这项研究的重要性可以从图 3-4 看出。传统经济理论预言，厂商将在价格P_e上生产Q_e。但是，成本加成定价至少会导致其他两个后果。第一，企业可能以平均总成本AQ_e及利润加价AB去生产Q_e。这将形成价格P_1，这里$P_1 < P_e$。第二，企业可能以平均总成本CQ_a及利润加价CD(假定等于AB)生产Q_a。这将形成价格P_2，而不是最优价格P_3。可以肯定，还存在其他可能性。

这种性质的第二项研究是理查德·莱斯特 (Richard

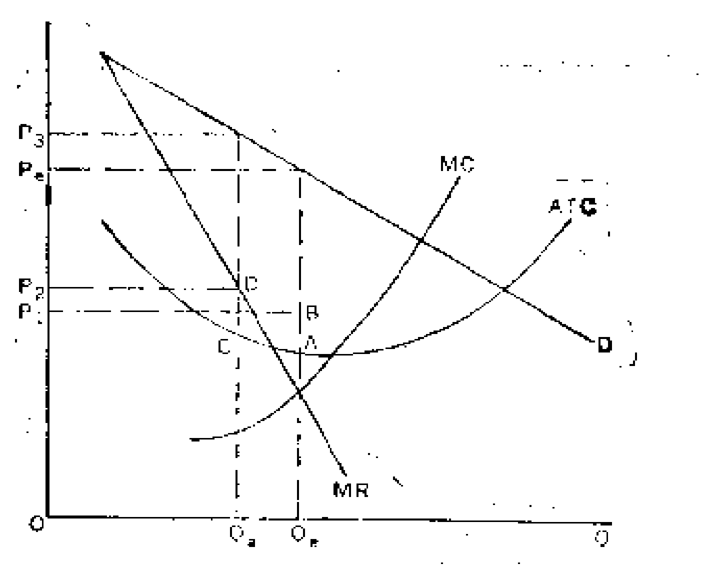


图3-4　成本加成定价与利润极大化

Lester, 1946年)做的。莱斯特将调查表寄给美国南部的430家企业，这些企业在美国南部的工厂中支付的工资，比它们在其他地方工厂支付的工资要低得多。莱斯特的资料以 68 份填写好的调查表为根据。对于是什么因素决定工厂的雇佣人数，28家企业的回答是只有产品的市场需求是至关重要的。在列举两种或两种以上决定因素的 28 家企业中，65％填的是产品的市场需求，9.7％填的是技术变化，9.5％填的是非劳动成本，5.1％填的是利润，只有 7.6％填的是工资和工资变动。这种情况令人惊奇，不仅是因为在经济学理论中强调的是劳动需求曲线，而且还因为在这 28 家企业中，劳动成本大约占去其总成本的 30％。在这项研究进行的时候，整个美国制造业企业的劳动成本约为总成本的 20％。因此人们料想，企业在进行雇工决策时，将对工资和工资变动十分敏感。但莱斯特研究的企业却不是这样。

莱斯特还询问这些企业家，当他们在南部的工人工资相对他们在北部的工人工资有所增加时，他们将如何反应。在回

答这个问题的 43 家企业中，占最大百分比(29.6%)的回答是说，他们将通过改进管理，改进全厂的生产或产品流程，改进工厂布局和进行激励试验来提高内部效率。只有 4 % 作出"正确的"回答：减少产量。莱斯特还报告说，他样本中的企业家一般都不考虑边际成本，他们普遍认为，估计边际成本或边际产量十分困难。

这里，人们再次面对看来与经济学理论相矛盾的行为。如果说企业不按经济理论描绘的方式营运，那么它们又是怎样运转的呢？它们实际使用的是什么营运方式？这对它们的绩效和稀缺资源的使用有什么影响？而且，在基尔比、罗斯塔斯、普拉顿和哈比森的研究中，他们所述的企业绩效似乎是低效率的，不过这种低效率并不是那种配置低效率。这就是说，不能因为这些企业比完全竞争的企业产量低，要价高，就认为它们缺乏效率。相反，应该把它们说成是内部低效率——成本比必要的高，产量却比可能的低。人们怎样才能解释这些资料呢？贯穿本书始终，我们将考虑各种不同的解释。在本章的余下部分，我们将叙述莱宾斯坦最初对X效率的论述。

3.3.4 X效率

为了解释这些资料，莱宾斯坦提出了X效率理论。顺着这个思路，我们考虑X效率与市场配置效率有何不同。这个问题如图3-5所示。一个完全竞争的厂商将按价格 P_c 生产 Q_c，而垄断厂商将按价格 P_m 生产 Q_m。面积 ABC 代表由于市场配置低效率而产生的全部福利损失。这就是说，面积ABC代表

由于厂商产量低于社会最优产出率 Q_c 而造成的福利损失。第1章把面积 P_mABP_c 解释为垄断利润——从买者转移到卖者的价值。这是因为，竞争厂商和垄断厂商的长期平均成本曲线和边际成本曲线被假定彼此完全一样，即如 $LRATC_c = LRMC_c$ 所示。这是假定厂商是成本极小化者的必然结果。如果厂商不一定是成本极小化者，那末就有充分的理由认为，产品市场竞争将给厂商施加较大的压力，使它们不论何时何地都尽可能地降低它们的成本。

可是，同样如图 3-5 所示，拥有市场力量的垄断厂商的成本曲线是 $LRATC_m = LRMC_m$。这就是说，对于既定的产出率来说，垄断厂商的成本总要高于完全竞争厂商的成本。这也意味着，图 3-5 中的面积 P_mABP_c 代表的是由于行业垄断而产生的较高生产成本，而不是垄断利润。它们代表垄断力量的成本，是一种低效率，但不是市场配置低效率。也就是说，这种类型的低效率并不涉及垄断厂商的产出率或要价。相反，它涉及的是与竞争厂商相比较而言的垄断厂商的相对

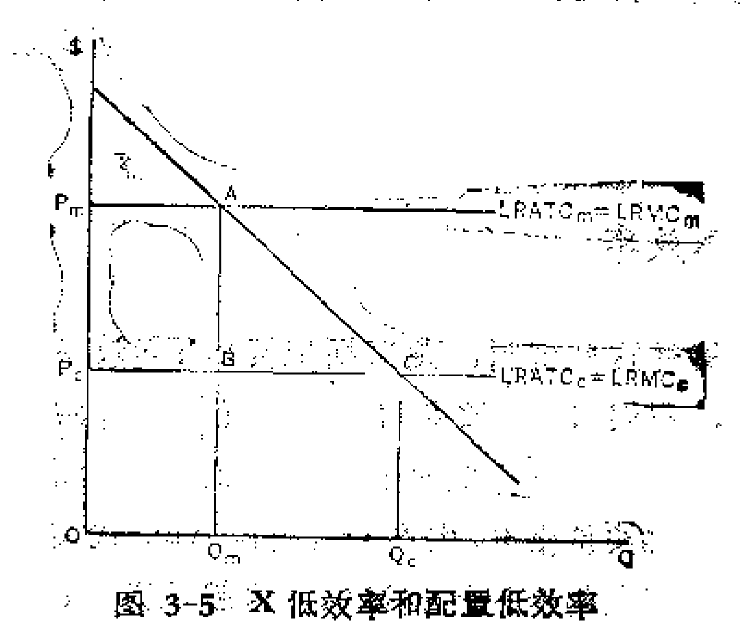


图 3-5 X 低效率和配置低效率

生产成本。这种低效率是一种非配置低效率，是一种未被当作经济学理论一部分的低效率，它的性质还没有完全肯定。因此，莱宾斯坦称它为 X 低效率。

支持这些非配置低效率存在的是些什么理论前提呢？也就是说，为什么既定的投入没有转化为预定的产出呢？莱宾斯坦讨论了四种因素。第一，劳动合同不完整。许多合同是完整的，即买者同意为某一商品或服务支付一定的金额，而卖者同意按双方约定的方式交付商品或服务。购买午餐、看电影或看垒球比赛是三个完全合同的例子。劳动却在非常重要的一个方面有所不同。尤其是，虽然劳动合同的支付部分讲得很明白，但是绩效方面却不很清楚。经理层可以用绝不含糊的词语宣布雇员每小时得到X美元，每周将得到X美元或每年得到X美元。但另一方面，他们却发现很难详细规定雇员应该做到的每一项活动的顺序与细节。支付火腿奶酪三明治是一回事，每天 8 小时、每年 50 周雇一个工人则是另一回事。火腿奶酪三明治可以立刻检验，一个雇员的持续绩效却很难检查，即使我们假定这样的检查是合乎需要的。

因此，工人可以决定自己的努力程度，这使得他们的绩效取决于他们的动力。莱宾斯坦补充说，"有许多东西得由习惯、权威、经理层可获得的某种激励方法、以及个人的处置和判断来解决。"与关于劳动合同的这种论述不同，新古典经济学隐含地将劳动合同与其他合同一样对待，并假定工作努力程度固定不变。

第二，生产函数不是完全确定或已知的。生产过程总是包含一种试验因素，所以企业并不总是预先就知道从既定投入和投入比率得到的产出数量。因此，一定的投入通常都有

许多不同的产出率。在图 3－6 中，我们把生产函数画成是一条粗带子，而不是一条细线。给定一单位资本和两个工人，产出可能在 Q_1 和 Q_2 之间摆动。第 2 章画出的生产函数像一条细线，就含有一定投入转化为预定极大产出的意思。

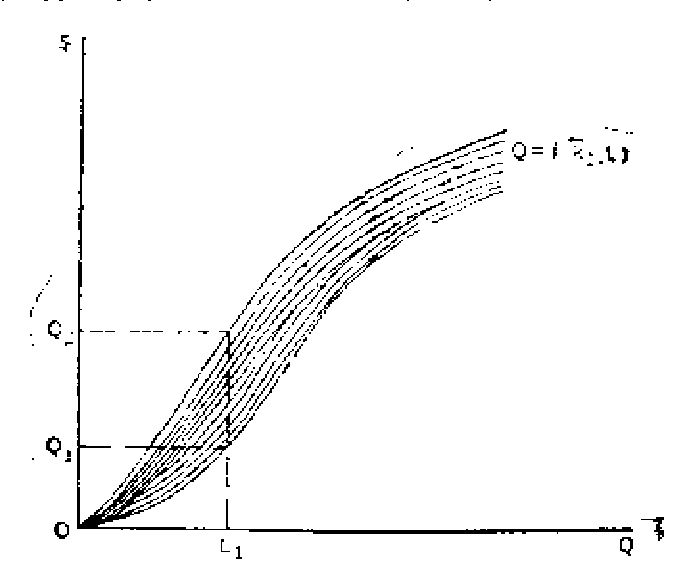


图 3-6　自由支配行为和"带状"的短期生产函数

第三，投入并不是都在市场上出售的，也不是所有购买者都能够平等地获得的。例如，在组织良好的市场里，管理知识常常不为人知，而且也不用来作交换。同时，管理的动机可能是一个正在按合适价格(即资金)"寻找"其他投入的厂商的决定性因素。知识和动机的作用对X效率的决定十分重要。标准微观经济学理论假定厂商实现了极大化，就是假定所有可获得的知识都已被充分利用了；而X效率理论则认为，在可获得的知识被使用的时候，动机是有影响的。下面，我们引用莱宾斯坦的一段话来说明这点。莱宾斯坦说："……就像资本和劳动可能利用不足一样，知识也可能没有被全部利用。更为重要的是，我们的许多知识还是模糊的。人们或许只能意识到它的存在，然而它却可能是关键的要素。假定

有充分的动机，人们是能够找到它的详尽性质并使它变得可供使用的。通常人们总是在脑力活动松弛的范围内工作。与利用不足的资本不同，知识是非常难以观察到的一个要素。"

第四，厂商可能彼此仿效而不是竞争。在这种环境里，没有理由相信厂商一定按它的生产和成本函数进行生产。莱宾斯坦说过："简单的事实是，个人和企业都不会尽力工作，也不会尽可能有效地搜寻信息。因为投入与产出之间的关系不是决定性的关系，所以，动机及其与努力和搜寻程度的密切关系的重要性就显现出来了。"

如果产出和成本不是完全预先确定的话，那末又是什么因素在决定这些变量呢？当然，技术是重要的，但是X效率理论引入一个称为"压力"的变量。压力指这样一种环境，在这种环境下，个人相对地感到"被驱使"去实现某种潜在的能力。在X效率理论中，据说压力是由竞争和逆境造成的，在这个理论最初形成的时候，成本被表明是由压力决定的。当然，这与说明成本是既定技术条件下的产量函数的标准公式不同，虽然在这个标准公式中，成本经常被假定是由技术决定的最小成本，但是在X效率理论中，成本却不被假定总是极小化的：有时候是，但并不总是。

图3-7表示压力对成本的影响。为此，莱宾斯坦假设，任何厂商的单位成本都受整个行业平均成本水平的影响。这意味着，每个厂商都以影响它自身成本的方式对全行业成本作出反应。特别是，当全行业单位成本较高时，每家厂商都能在保持竞争力的同时允许它的成本上升。另一方面，当全行业单位成本下降时，为了保持竞争力，每家厂商都将设法降低它本身的单位成本。在图3-7中，我们假设前一时期(t—1)

的全行业成本有助于厂商形成对现期(t)全行业成本的预期，并由此影响现期的厂商单位成本。45°线是所有(t－1)时期的全行业单位成本等于(t)时期厂商单位成本的点的轨迹。换言之，厂商会对全行业成本按比例作调整，使得两者彼此相等。

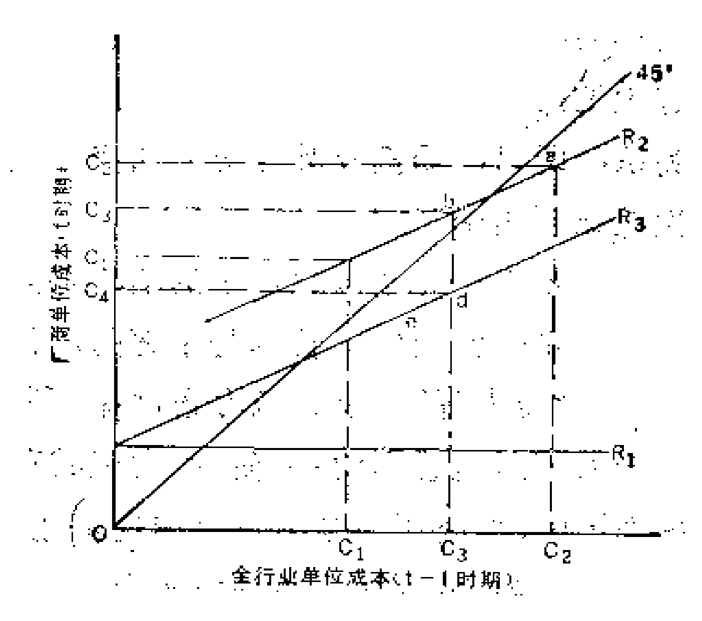


图 3-7 压力、技术变化和成本

R_1表示成本极小化者的"反应曲线"。这种厂商成本对上述的行业压力不作反应。另一方面，R_2是那些其成本对压力作出反应厂商的反应曲线。例如，如果厂商预期全行业单位成本是C_1，那末它的成本就将是c_1。如果厂商预期全行业的单位成本是C_2，那么它的成本就将是c_2。反应曲线的精确图像当然不像我们这里所画的样子。但重要的是，厂商成本受其他各个厂商绩效的影响。

加上反应曲线R_3可以显示X效率增加和知识增加所造成的成本下降。假定我们从全行业成本为C_2，厂商成本为c_2的点a开始。如果全行业成本下降到C_3，那么厂商成本就将下

降到 c_3，即从点 a 沿曲线R_2移动到点 b。如果知识增加，那就可以看到整个曲线向下移至 R_3。随着厂商从点 b 移至点 d，这时厂商成本将降至 c_4。如果全行业成本再次下降，那么厂商成本就也会再次下降，沿着 R_3 移到例如点 e 处。不管实际过程如何，厂商成本不再假设是预先确定的，它受降低成本动机的影响，当然也受知识（技术）的影响。因此我们不再把成本曲线（像第 2 章图 2－4 所示那样）画成细线，而是如图 3－8 所示，画成一条宽带。在图 3－8 中，ATC_1 表示任何产出水平能导致一个根据压力和动机变化的成本水平。例如，Q_1 将在 C_1C_2 间进行生产，确切的单位成本则由降低成本的压力与动机决定。这个间距相当于图 3－7 中由于压力变化从点 a 到点 b 的移动。另一方面，知识的增加将使整个ATC曲线向下移到 ATC_2。现在产出 Q_1 能以单位成本 C_3 生产，但在 C_3C_4 间的那个范围内的生产，将由降低成本的压力决定。ATC曲线向下移动，类似于图 3-7 中反应曲线从 R_2 向下移动到

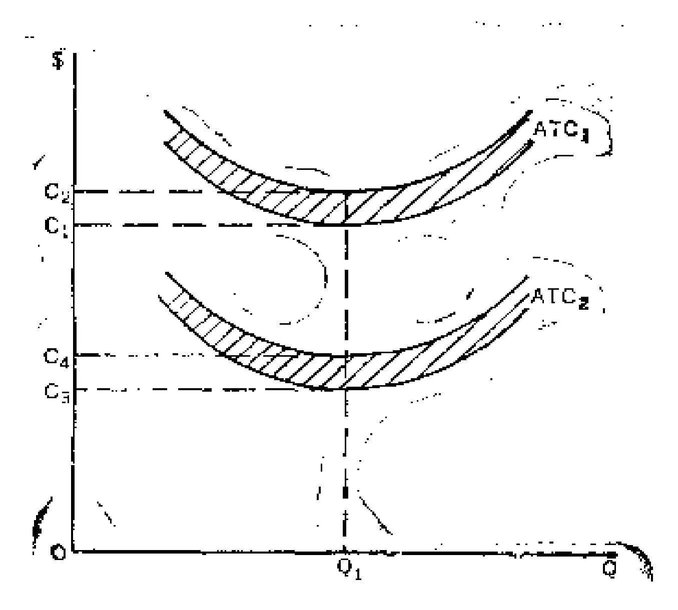


图 3-8 X低效率和技术变化对成本的影响

R_3以及随后从点 d 到点 e 的移动。莱宾斯坦概括这些论点说，"……由于各种各样的理由，个人和组织通常不会像他们可能做到的那样努力和有效率。在竞争压力较轻的环境里，许多人将以花费更多的努力、搜寻以及监督他人活动带来的负效用来换取不感到压力和较好的人际关系所获得的效用。但是在竞争压力较大因而进行这种交换的代价也较大的环境里，他们将较少用努力的负效用来换取摆脱压力的效用，如此等等。"随着我们进一步的阐述，我们会看到，这种效用极大化努力的决策思想将被非极大化行为的可能性取代。

3.4 结论

20世纪 60 年代早期，经济学家提出了几种非利润极大化的厂商行为理论。那些我们称之为"复合目标函数"的理论假设存在极大化行为，但不是利润极大化行为。在考察非极大化行为方面，莱宾斯坦的 X 效率理论前进了一步。以后两章中我们将论述 X 效率理论。

4 X效率理论 1

4.1 导言

在上一章中，我们考察了与正统新古典理论(似乎)不一致的资料，以及莱宾斯坦解释这些资料的最初尝试。在本章和下章中，我们将更详尽地分析他的解释——X效率理论①。

X效率中的X代表造成非配置型低效率的未知因素。这就是说，上一章所述的资料，显示了增加企业内部生产率，从而降低企业成本的未被利用的低成本机会。这些(X型的)未被利用的机会，不应该与规模经济或范围经济相混淆。这些未被利用的机会，是由于诸如动机或缺少动机这样的因素，由于"普通常见的"人的惰性，和由于并不总是旨在实现成本极小化的行为造成的。这些因素，在使成本超过极小化水平时，被认为是低效率的一种形式，但不是配置的——价格或市场——低效率。这些低效率与价格和市场无关。它们与影

① 第4章(和第5章)是X效率理论重要论点的一个概述。对此富有兴趣的读者，可查阅本书参考书目中列举的莱宾斯坦的著作。

响企业绩效（相对企业的潜力而言）的企业内部活动有关。通常，X低效率总是就企业的产出率或成本而论的，但是，它也可以就企业的技术适应率来加以讨论，本书不打算这样做①。

在进入这个理论之前，我请读者把注意力放回到第一章导论中的表1上面。表1旨在说明，X效率理论考虑到了个人会出现类型1误差和类型2误差这样两种可能性。另一方面，我认为新古典理论是假定不存在类型1误差的，并试图说明为什么被解释成类型2误差的东西，事实上根本就不是什么误差。在本章和下一章考察的X效率理论中，读者应该弄清楚，X效率理论是如何解释类型1和类型2这两种误差的可能性的。

读者可能还回忆得起，企业行为的复合目标函数理论假定了极大化行为，而莱宾斯坦的X效率理论，如同我们将表明的那样，既考虑到极大化行为，也考虑到非极大化行为。因此，本章和下章的任务之一，就是说明个人的"内部"（个性）环境和"外部"环境如何能产生这两种类型的行为。

4.1.1 新古典理论和X效率理论的假设

要解释这些资料和非配置（低）效率，部分工作至少是要改变新古典理论的某些假设，并探究它们的含义。现在，我们来陈述莱宾斯坦提出疑问的新古典理论假设和他自己的理论假设。受到怀疑的新古典理论包含这样一些假设。第一，

① 莱宾斯坦（1978年a）在他论经济发展和X效率理论的著作中讨论过技术适应问题。

企业和家庭是恰当的研究单位。第二，个人行为以理性为特征，就是说，只要是人的行为就是实现极大化的行为。对于企业来说，这意味着利润极大化(成本极小化)；对于家庭来说，这意味着效用极大化。第三，企业被看作好像只追求单一目标的极大化。例如，如果这个目标是利润，那么第三个假设的意思就是，雇员(代理人)——经理和其他人——的行为就是为了使企业主(委托人)的利润实现极大化。第四，与理性假设有关，人的行为肯定对环境变化作出充分反应。第五，劳动合同是完整的。第六，假设人的努力程度既定。

另一方面，X效率理论作了如下假设。第一，只有个人有思想和行动，因而个人才是恰当的研究单位。研究由个人入手，从个人行为进到集体(groups)行为。第二，假设个人行为既包含理性因素，又包含非理性因素。X效率理论并不否认理性行为，注意这一点对于弄懂X效率理论很重要。但是，X效率理论假设，人并不总是表现出(完全)理性。因此，在一般情况下，只假设个人具有选择的理性，而不是完全的理性。第三，企业主的利益与其代理人的利益并不总是一致的。第四，人的行为并不总对环境变化作出反应。相反，人的行为有时具有深受习惯影响的惰性特征。因此，有人说个人经常在"惰性区域"工作。第五，假设劳动合同是不完整的。第六，假设五意味着，努力程度是一个任意的变量，而不是常量。最后，X效率理论最适合于不完全市场。

新古典对理性的假设就这样被应用，使得生产过程好像就是一个机械过程，企业好像是一部把投入变成最大产出的有效转换器。另一方面，X效率理论的假设考虑到不确定性

(indeterminacy)。我们将按如下顺序论述X效率理论。第一，我们讨论个人的心理和行为。第二，讨论集体的行为。第三，讨论环境对X效率的影响。第四，讨论这个理论的涵义。第五，把这个理论应用于生产率决定。

4.2 X效率理论中的个人

关于X效率理论的心理方面，我们分两部分来讨论：第一部分，是作为个人的个人；第二部分，是作为集体成员的个人。第二部分——作为集体成员的个人——把X效率理论的心理因素和环境因素联接了起来。

X效率理论主要是针对两件事作出的反应：(1)把经济事务——例如生产过程——描写得过分僵化或全由技术决定的理论；(2)看来存在与这个理论不符的正式和传闻的证据——其中有一些已在第二章提到。如果生产不仅仅是技术力量的结果，那么决定投入-产出关系的又是什么呢？如果这种证据事实上需要解释的话，我们又该怎样解释它呢？部分来说，X效率理论是通过考察某些通常使用的词汇和术语的构成和含义，包括决策、努力程度（体力或脑力的）、压力、极大化和理性，来说明这两个问题的。在每一种情况下，读者都能看到这样一个倾向，即认为每一个这样的词语都应该由一个可能性区域来定义。这就是说，决策可能是"严格的"，也可能是"松散的"，努力程度可能从少量到大量不等，压力也可能这样。读者还能看到，X效率理论有研究这些词语的行为和程序方面的倾向。例如，努力的组成部分是什么？在什么条件下更可能出现较高的努力水平？或者像赫伯特·西

蒙① (Herbert Simon) 一样，在程序意义上使用理性。也就是说，要作出理性决定，一步一步地需要哪些程序?

4.2.1 双重人格

个人能对X(低)效率作出贡献的观念，来自这样一个推测的可能性，即个人的行为受两种表面看来相冲突的倾向的影响，每一种倾向都激发一种不同类型的行为模式或对于环境的反应。一种倾向是确立或坚持标准，努力追求极大化，这种努力是经过计算的和注意细节的。换言之，这方面的人格是我们"理性的"自我，是在决策前使用"严格的"决策程序的那部分自我，是(在某些时候)把注意力集中在解答的细节上的那部分自我。莱宾斯坦称我们人格的这个方面为"超我功能"(superego function)。(虽然弗洛伊德曾使用过这个术语，但认为这里使用它就暗示莱宾斯坦试图把弗洛伊德的心理学引入经济学理论，那就错了。)"约束"我们的就是这个超我功能，这就是说，我们的这一部分是理性的(就上面使用的意义而言)，我们的这一部分将努力坚持高标准而不是低标准。

我们每个人都有的另一个倾向是"倒退回去"，使用"松散的"决策程序，遵循我们的"动物精神"。莱宾斯坦称这种倾向为"本我功能"(id function)。就是这个本我功能导致我们"不受约束"，就是说，使我们不愿意(但不一定不能够)去计算，去注意细节和采取理性行为。

---

① 关于程序的和实体的理性的论述，可参见西蒙(1957年，1959年，1978年)的著作。西蒙理论和X效率理论之间的某些异同点，莱宾斯坦(1976年，1979年，1985年)的著作里有论述。

X效率理论假定，一般说来，这两种功能会以导致两者相互妥协的方式影响每个人。这就是说，我们每个人都会在我们觉得必须这样行为的方式(超我功能)，和我们喜欢但与履行职责或一系列标准（不管这些标准是由内部还是由外部施加的)不符的方式(本我功能)之间达成一种妥协。换言之，每个人都会形成一种使他在心理上感到“舒适”的妥协。

受约束与不受约束这两个词，一直在个人能够并愿意注意细节、进行计算和合乎理性的情况下使用。现在，受约束和不受约束越来越被看作是能被连续尺度衡量的人格特征的两个“极端点”。这就是说，人格在它的一端表现为“完全关心约束”，在它的另一端表现为完全不关心约束。因此，一个个人可以运用他或她的全部力量去注意细节，计算成本和利润。这样的个人就是“完全或充分理性的”个人。另一方面，可以想象一个个人把少于100%的力量放在作出信息充分的决策所必须的细节和计算上。这样的个人可以称为“有选择的理性的”个人。因此，理性是一个表现出“完全关心约束”的“经济人”的连续变量。因此，经济人是一个“极端情形”，是某些人在某些时候采取决策程序的特征，但却不一定是所有人在所有时候的特征。

4.2.2 有选择的理性

现在我们要进一步研究与充分的和有选择的理性(selective rationality)相一致的行为或程序因素。X效率理论中这个部分的目的在于发展与充分的以及有选择的理性尽可能一致的决策程序一览表。这样的程序一览表使理性概念具有行

为的意义，因此否定了理性是任何人在任何时刻的行为特征的简单断言。

我们要探究理性决策的六个要素。它们是：(1)现实地评估环境；(2)非反射性地评估环境；(3)独立判断；(4)高度的敏感性；(5)不拖延决策和行动；(6)从经验中学习。这些要素中有些可能彼此重叠。例如，一个非反射性（非机械反射）评估可能隐含一个较为现实的评估。虽然这会减少要素的数目，但不会消除要素本身。而且，这里我们主要关心的是建立可以研究理性的框架或步骤。为此，理性概念必须是可以操作的。这个一览表就是一种可行的办法。

这个一览表中的每一个要素都被看作是连续变量。例如，一个决策可以完全根据对环境的现实评估，也可以完全根据所希望的想法。另一种可能是，一个决策还可以既根据对环境某些部分的现实评估，又根据对其他部分所希望的想法。对一个项目风险的评估会使现实主义的重要性变得十分明显。

对环境的反射性评估是一种机械反应或决策。它是以决策者体验到的第一次冲动为根据的无信息的决策。另一方面，非反射性评估的特征则是"冷静地"评价有关的事实，机械反射不是自发的同义词；自发指的是一个专家几乎在瞬刻间非反射地、"无意识地"作出的决定。

判断的独立性可以有从完全独立到完全不独立的区别。不独立至少包括下列决策程序：从那些掌握信息比自己还少的人那里接受劝告；行动完全是为了让他人高兴或得到他人的允许；无关紧要的仿效。和理性的其他要素一样，独立与不独立常常是一个程度问题，而不是一个有或无的问题。

完全关心约束——完全理性——被看成是对环境变化敏

感的决策的结果。只有那些高度敏感的人才能利用有关环境资料如价格、收入、需求等的细微变化来获得好处，正是这些资料保证他们自己的行为会对环境作出合宜的反应。

今天能做的事决不推迟到明天去做，尤其是当你打赌明天不会再来的时候。（永远如此！）时间拖延指的就是上面这种推迟，以及希望明天会有能力和（或）意愿变得更加合乎理性，因而今天就可以实行松散决策程序的做法。时间拖延明显地不同于等待完成信息搜寻与分析的那种决策的推迟。

决策的结果依赖它的信息内容。有些决策依据的是所有过去的有关经验，而不问它们是“成功的”经验还是“失败的”经验。另外一些时候，我们可能只想起成功的经验，并根据它们相应地作出当前的决策。如果其他情况保持不变，决策的结果将根据这种“有选择的学习”的程度或我们如何向过去的经验学习而有所不同。总之，理性的成分被认为是连续的，而不是二元的（有或无）变量。任何个人在任何时候表现出的理性程度，都是超我功能与本我功能对个人行为双重影响的结果。而且，完全理性被看作这样一种决策程序的结果，其特征是：最优地利用了现实性、非反射性、独立性、高度敏感性、没有时间拖延和从全部（有关的）过去经验中学习的意愿和能力①。

4.2.3 有选择的理性和启发式决策

启发式决策常常被说成是一种既能减少复杂性又能不失

① 见弗朗茨（1980年）关于莱宾斯坦的（有选择）理性概念和心理学家亚伯拉罕·马斯洛（Abraham Maslow）的概念之间的相似性的论述。

精确性的办法。购买和你邻居一样的商标和质量的商品，是经常使用的一种启发式方法。那些使用这种方法的人相信，给定"正确性"和"良好鉴赏力"，这些方法可以使他们减少搜寻的需要。凭学历甄别潜在的雇员，是普遍使用的启发式方法。可是，它虽然有用，却也会导致严重的判断失误。有些这样的启发式方法与有选择理性的要素相一致。阿莫斯·特维斯基和丹尼尔·卡内曼(Amos Tversky and Daniel Kahneman,1974 年)提供了三种这类启发式方法。它们被称为：(1)代表性方法；(2)可得性方法；(3)固定性方法。

用代表性方法作决策时，个人要决定A怎样成为B的代表。特维斯基和卡内曼用估计一个特定个人在一项特定职业中工作的概率所使用的程序来说明代表性。这种启发式方法引导人们通过考虑个人对每种职业而言具有怎样的代表性这个问题，来估计这些概率。如果告诉人们，阿莫斯有健壮的肌肉、秃顶、大胡子、头上刺有花纹而且经常尖叫，那么启发式方法就会使人们认为他极有可能是一个摔跤手，而不是精神病医生。当然，代表性不受影响概率判断的因素的影响。一种这样的因素就是先验概率——例如，从事摔跤和从事精神病医生职业的人口百分比。特维斯基和卡内曼报告说，先验概率是在没有个性特写的情况下被正确采用的。就是说，只有在不能判断代表性时，才使用先验概率。

可得性方法也是一种启发式方法。靠这种方法，人们可以根据记起一个事件的容易程度来判断该事件的概率。由于经常发生事件的实例比不常发生事件的实例容易回忆，所以可得性启发式方法看来是很有用的。但是，根据特维斯基和卡内曼的说法，由于可得性还受到除频率和概率外因素的影

响，所以它也会导致判断错误。例如，容易回忆起的事例被判断为大量发生的事例。在一次研究中发现，当一张名单上较少出现的性别是由较著名的人物组成时，受试者对这张名单上是男性居多还是女性居多的判断就会出错。可得性还受到比较近期的事件的影响。例如，在目睹一场车祸后，车祸的主观概率就提高了。特点也会影响可得性。如果有人看到汽车起火，然后又在报上读到此事的报道，那么汽车起火的主观概率就会增大。特维斯基和卡内曼还引述了许多别的例子。

固定性出现在从某个初始值开始的最后估计偏向这个初始值的时候。因为不同的起点（它能在问题的表述中获得或者从初始的局部计算中推出）会产生不同的估计值，所以固定性也会导致判断错误。在特维斯基和卡内曼报告的一次研究中，要求两个受试组估计非洲国家在联合国中占有的百分比。一组获得的起点值是10%，而另一组获得的起点值是65%。这两个受试组的中间估计值分别为25%和45%。

在所有这三种启发式的场合，偏差的发生都并非归因于像希望的想法这样的动机效应或奖罚造成的判断扭曲。应该说，它们是起因于启发式方法的运用本身。此外据报道，外行和有经验的研究人员都会犯这种判断错误，尽管有每天的和日积月累的实际生活经验，这种错误也会继续存在。特维斯基和卡内曼得出结论说，判断是能够得到改进的。舍马克(Schoemaker,1982年)在他评论和分析预期效用(EU)模型的著述中报道了相同的行为。他的一个结论是，预期效用理论（在这个理论中，个人被假定是使他们的预期效用实现最大化）不论是作为一种描述性的理论，还是作为一种预测性的理

论，它都是失败的，因为它没有考虑到人类认识简化的倾向。许多研究，包括特维斯基和卡内曼的研究他都作了评论。

4.2.4 有选择的理性和工作绩效

考察一下表示理性和压力程度之间可能替代的“无差异曲线”。这里使用的压力一词，表达的意思是：由个人自身内部机制，或者由某种外部力量，施加在个人身上的旨在改变他或她行为的“力量”。很清楚，压力会表现出不同的强度，从温和的责备到警告，直至诉诸法律的行动。图 4-1 所示的两组无差异曲线，意在说明理性和压力（理性不存在时，压力就是理性）是可以相互替代的“状态”，但是决没有假定它们是可以完全替代的状态。这就是说，对于既定的效用水平而言，如果理性水平事实上降低的话，那么个人愿意经受更大的压力就是理性的。相反，如果效用不变，那么和更高的理性相伴随，压力水平降低才是理性行为。无差异曲线的斜率指出，例如现在为了减少一点理性，个人愿意忍受多大的压力。

成为理性的压力可以是预期的压力。在这种情况下，图 4-1 中无差异曲线表示的是，现在显示的理性水平与未来某个时候为保持这种理性水平所需的预期的未来压力水平之间的替代关系。虽然我们的无差异曲线假定，现在感觉到的压力与预期的未来压力对行为有同样的影响，但我们却没有理由假定情况总是这样。

图 4-1 中标上 I_5 的无差异曲线称为“本我无差异”曲线。本我无差异曲线凹向原点，这表明理性对压力的边际替代率（MRS）是递增的。这种情形意味着个人既不喜欢太多的理性，

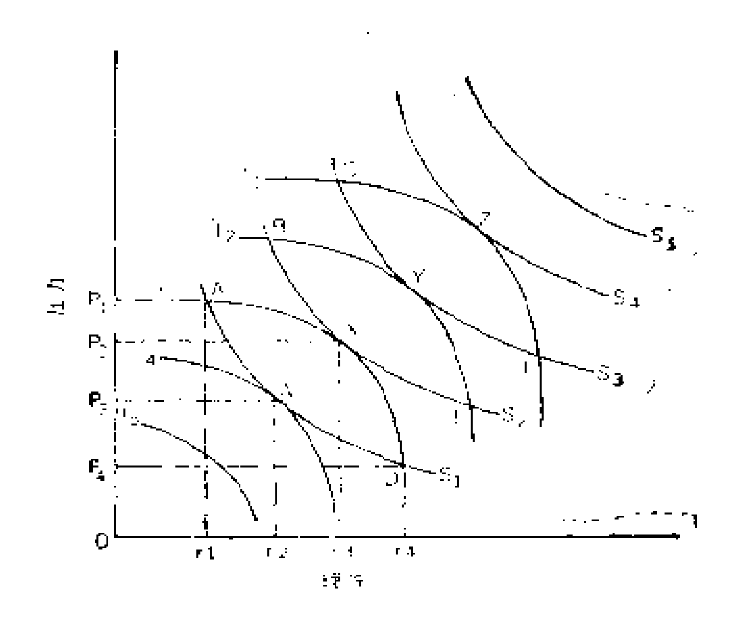


图 4-1　人格、有选择的理性和X低效率

也不喜欢太多的压力，因此，压力(理性)量的增加，每次都会以牺牲更多的理性(压力)作为代价①。在图 4-1 中，最接近原点的本我曲线代表较高的效用水平。就是说，本我功能的影响是，如果其他情况不变，一个人宁愿按仿佛他(或她)没有责任、义务和标准这些观念那样行事。可见，较少感到压力和较少关心现在的约束会增加个人的满足。

图 4-1 中标上S_1的曲线是"超我无差异"曲线。超我无差异曲线凸向原点，表明理性对压力的边际替代率递减。这意味着个人宁愿坚持标准等等，因此，压力(理性)量的减少，每次都会以牺牲更多的理性(压力)作为代价。在图 4-1 中，离原点较远的超我无差异曲线代表较高的标准。因为，$S_3>S_2>S_1$，超我无差异曲线集合表明较高的标准将带来较大的

① 这里使用无差异曲线并不意味着个人在作极大化的决策。就是说，他们并不根据极大化行为规则在边际上进行替代。相反，这些曲线仅仅表示个人至少具有有选择的理性。我愿意使用满足这个词，而不愿意使用效用这个词，因为后一个词隐含着充分的理性行为。

满足。

双重人格对效用的这种影响，可以借助图 4–1 作进一步的说明。我对图 4–1 的解释在若干重要方面不同于莱宾斯坦的解释。先考虑点A。在点A，个人按相对低的标准(S_1)工作，体验压力水平 p_1 和理性水平 r_1。按这样的低标准工作，个人通过超我机制感受到相对小的满足，但通过本我机制却感受到相对大的满足。点A不是一个最舒适的位置。为理解这一点，只要想象一下这个人用现在的理性去替代(预期的未来）压力。图 4–1 允许他有两种选择。第一，通过把他的理性(经由理性的任何一个或几个要素)增加到 r_3 并把他的压力减小到 p_2，他能够按较高的标准(S_2)工作，而他不会损失本我舒适。这个选择的结果使他处于点X上，这是一个新的工作模式：产生更多的超我满足，没有本我舒适的损失。他继续留在 I_3 上，同时把他的位置从 S_1 移到 S_2。因此，点A是一个轻松有余的工作模式。同时，它也是一个不舒适有余的工作模式！因此，可以把点A看作是对个人要求不充分的工作的结果，而作为结果，这个工作是不舒适的。

这个人面临的第二种选择是按同一标准(S_1)工作，同时增加他的本我舒适。把理性增加到 r_2，同时把压力减低到 p_3，使位置从A移到W，就可以做到这一点。在点W上，他的本我舒适从 I_3 增加到 I_4，而他的超我继续在 S_1 水平上得到满足。同从点A移到点X相比，从点A移到点W包含理性的较小增加和压力的较大减小。因为，向点W的这种转移有利于本我而又无损于超我，所以，压力可望有相对大的减小。

因此，点A不是最优舒适的工作模式，理由有二。第一，因为如上所述，它不必要地令人厌烦。第二，因为倘若个人

继续从事这项令人厌烦的工作，他就应该利用这项工作，在工作中相对小地增加理性(从 r_1 增到 r_2)，同时相对大地减小压力(从 p_1 减到 p_3)。换句话说，这位“过分能干的”个人只要稍稍多关心一下这项工作，就能从这项令人厌烦的工作中得到更多的满足。由此产生的压力损失补偿他绰绰有余，足以使他处于较高的总满足水平上。这些解释也同样适用于点 B 和点 C，而点 Y 和点 Z 则代表最佳工作模式或最舒适水平。

现在考虑点D。点D代表人格“过度紧张”的个人。和点A一样，在点D个人按相对低的标准(S_1)工作，但是，现在他却处在较高的本我舒适水平(I_3)上。在点D，个人具有理性水平 r_4 和压力水平 p_4。像点A一样，点D也不是最舒适的位置。如同在点A上一样，在点D上，个人也有两种选择。第一，移到点W，就是说保持同样的超我标准(S_1)，但将本我舒适增加到 I_4。因此在点D上，对于既定的超我标准，个人没有使他们的本我舒适达到极大，在这个意义上个人是过度紧张的。对于他所保持的超我标准来说，他是过度理性了。他太努力了——没有必要。他正使自己超越必要的界限。因此，从本我舒适的观点看，个人可能是“过度”理性了。

为了移到点W，个人应将理性从 r_4 减小到 r_2，同时将压力从 p_4 增加到 p_3。点D不是最舒适点，也不是点A那样令人厌烦的点。在点D，个人的过分理性，对现在细节的过分关心，简直就是“精力”的低效率，他正在“做劳而无功的事情”。他需要“轻松愉快”。点E和点F表示相类似的状况。

第二个选择是把超我标准增加到 S_2，同时又不损失本我舒适。移动到点X就可以实现这点。这样的移动要求理性从 r_4 减小到 r_3，同时压力从 p_4 增加到 p_2。同点X相比，处于点

D的个人在就他所享受的本我舒适水平而言他是过度理性的意义上，他是过度紧张的。这种个人做着"劳而无功的事情"，把较多的注意力放在保持必要的超我标准上。这样，他就很可能"只见树木不见森林"。他能以较小的关心程度，保持较高的超我标准，从而通过他的超我得到较多的满足，而又不损失本我舒适。同移动到点X相比，移动到点W要求大量减小理性而稍微增加压力。因为移动到点W将增加本我舒适，所以，这些相对的数量不会令人吃惊。虽然，具有点D所示的那种工作模式的个人能够通过努力达到较高的标准，而又不损失本我舒适，但这并不意味着点D是令人厌烦的（如点A那样）。在点D上，个人并非厌烦，相反他过分关心现在的细节，导致"精力"的低效率，他做的是"劳而无功的事情"。因此，点D上的个人是过度理性的。他需要轻松愉快。在图4-1中，点E与点F和点D相类似。

因此，从超我满足的观点来看，个人是能够过度理性的。就取得较高超我标准意味较小X低效率而言，在某些情况下，减少理性就能减小X低效率！

我们引入了"过度理性"这个概念。据说理性概念及其要素是连续变量。图4-1告诉我们，"完全"理性不需要与这个连续尺度的二个极端相联系。假如我们把个人的理性确定为值域从1到10的连续变量，那末完全理性——由一个特定的个人，在特定的时刻用最具有X效率的手段完成特定功能所必须的理性水平——不需要与10这个值相联系。在点D、点E或点F的情况下，完全理性的值就被认为小于10。

在图4-1中，直线WXYZ代表最优舒适，在这条线上，不论是本我满足还是超我满足的增加都意味着另一方的减

少。因而，点A、B或C不是最优的，这些工作模式的结果是过分厌烦和不舒适。更多的理性将增加满足程度。另一方面，点D、E和F代表个人会过分不舒适的工作模式，因为它们是过分理性即过分关注细节了。因此，最优舒适要求人格、工作条件及理性表现程度的恰当配合。点W、X、Y、Z就代表这种最优配合。

另一方面，读者应该了解，图 4-1 中的任何一点，不论是不是在"契约曲线"WXYZ上，都是(体力上或精神上)舒适的点。在契约曲线上的点并不意味着X低效率为零；从人格观点看，它们是最优舒适的点。例如，在点D上，个人通过将他的理性从 r_4 减少到 r_3，可以达到较高的超我标准。但是，在点X上的超我标准 S_2，并不一定与他工作绩效中的较高X效率相联系。标准S_2超过S_1，但是他的特殊工作在绩效不存在任何X低效率之前就可能要求达到 S_5。因此，点X(所有其他在或不在契约曲线上的点也一样)仅仅是最优舒适点，而不一定是X低效率为零的点。

可是，我们所讨论的图 4-1 之所以这样画法，是为了使个人总是能够找到最优舒适水平。例如，在点A或点D上的个人能够找到由提供最优舒适的点W或点X代表的工作。同样，在点B或点E上的个人能够在由点X或点Y代表的工作中找到最优舒适。点C和点F代表可以转变为最优舒适点Y或点Z的工作模式。但是，考虑一下超我曲线 S_5 和本我曲线 I_5。超我曲线 S_5 不与任何本我曲线相交。这意味着标准S_5将导致不舒适，因为对于他既定的人格来说，这是一个太高的难以维持的标准。尤其是，对于他既定的本我来说，这更是一个太高的难以维持的标准。如果 S_5 代表一个特定的工作或

职业，那末我们就可以预期此人将寻找别的工作以达到最优舒适水平。如果零X低效率要求达到标准S_5，那末只要其他条件不变，零X低效率就不可能达到。

另一方面，本我曲线I_5不与任何超我曲线相交。这意味着，虽然工作模式I_5代表相对大量的本我舒适，但它还是不舒适的，因为这人无法按他为自己规定的，代表一个可接受自我形象的标准生活。产生本我舒适I_5的工作，对这人来说是不能接受的。再者，人格追求舒适，但不一定追求零X低效率。而且，虽然我们可以预期个人将趋向契约曲线上的各点，但是我们没有理由相信，这样的一种运动会特别迅速或精确。正如我们将要论述的，X效率理论假设，特定的行为模式一旦形成，就很不容易改变。

4.3 个人努力和惰性区域

企业不会自动地将投入转变为技术最优的产出率。这有几个理由，其中包括这个事实，即企业购入的是劳动时间，而生产需要的却是劳动努力，也就是莱宾斯坦称作“按照或超过某个技能水平的”直接努力①。这就是说，人力投入的最重要因素——直接努力——不是直接购入的，这是人力投入和非人力投入的许多区别之一。因此，可以把努力看成一个人对他（或她）自己的精神和（或）由外部环境确定的动机作出反应的结果。

这里使用的努力一词既包括它的体力方面，也包括它的脑力方面。这就是，体力努力指的是行走、提物、说话等等

① 此话引自莱宾斯坦（1976年）。

形式的体力运用。脑力努力指的是脑力的运用。这两种努力都可以有不同程度的表现，因此都是变量。除了解释这种变动性的生理原因以外，工作环境也会提供它自己的刺激。

这种工作环境的一个重大要素就是"不完全"的雇佣合同。由于雇佣合同不完全，所以它与非雇佣合同，例如与饭店的合同就有所区别。从熟食店叫一份面包圈和奶油乳酪，就代表一个"完全的"合同，也就是说，你要一份面包圈和奶油乳酪，并同意支付一定数量的货币。你确切知道你要的（买的）是什么，企业也确切知道销售以后他们将得到的是什么。但是在雇佣合同的情况下，企业购买的是时间，但生产使用的却是雇员的直接努力。这种安排或合同是不完全的，因为在企业购买的时间和它使用的努力之间没有固定的替代关系。熟食店同意并且事实上将一份面包圈和奶油乳酪放在了你的面前，你也同意并且事实上支付了确定数量的货币。可是在雇佣合同中，同意的事情和做到的事情却不一定相同。换言之，大部分工作中的多数雇员都有一定的自由处置余地。这里我们将指出努力的要素，以及雇员自由处置的方式，但是不作进一步讨论。

据说，努力有四个要素：(1) 构成工作的活动(A)；(2) 完成工作的进度(P)；(3) 产品的质量(Q)；(4) 工作的时间模式(T)。可能还有其他要素，上述要素有些也可能彼此重叠。这些要素构成一张也许并不完全的清单。但是，为了使努力的概念变得可以操作，这样一份清单是必要的。什么活动构成一项特定的工作？什么活动会增加做好这项工作的可能性？绝大多数个人在开始做一项工作时是拥有一些信息的。但是他们却绝不可能获得关于这项工作各方面细节的确切说明。

因此，个人可以作出关于如何完成这项工作的经过计算的(完全理性的)决策，也可以只是按某种方式开始和继续进行工作，他(她)可以仿效他人，也可以使用许多其他可能的决策技术中的任何一种。虽然有点犹豫，但为了简化起见，我还是把这些决策技术的使用视为"选择"的结果。因此，个人必须以某种方式选择或解释这种工作。在这种情况下，他(她)必须至少就这项工作的某些方面，对他(她)要做什么作出选择。他(她)还必须至少就某些这样的方面，对他(她)完成工作的进度作出选择。关于质量采取某些自由处置办法也是可能的，在这里，关于质量的选择与个人在工作中注意力的集中程度有关。最后，他(她)也能够对时间模式或节奏作出选择，这包括一天的工作时间或"做做停停"(stop go)模式。

一旦作出选择，这四点决策就构成工人的"APQT束"。在个人选择这种努力水平的范围内而言，他(她)将选择其APQT束。否则，APQT束就由企业"预先确定"(pre-set)。至此我们一直认为，至少部分地存在自由选择的APQT束。自由选择的APQT束会有多少普遍性？我们认为，它们具有许多合意的特性，因此很可能具有普遍性。自由选择的APQT束对努力和企业的X效率有什么影响？关键在于，努力成为一种变量，它可以产生也可以不产生理性行为，即成本极小化和完全的X效率。

预先确定APQT束对企业提出了几个问题。第一，为全体非经理雇员预先确定APQT束要求大量知识，从而要求做大量的搜寻工作。因此，预先确定APQT束的搜寻成本可能很高。第二，如果由经理人员预先确定APQT束，那么很可

能，他们需要确信非经理人员会根据他们工作。因此，预先确定 APQT 束的监督成本也可能很高。第三，因为企业可能根本不相信你会做出如何最好完成你工作的决策，在这种企业工作可能产生精神成本，导致旷工和怠工。第四，预先确定 APQT 束不能利用非经理雇员在经验中积累的知识，这就使企业失去了一个重要的人力资本来源。第五，“死扣规章进行怠工”常常被看成是故意破坏一个企业的最厉害的办法，因为它使企业在不断改变的环境中所必须的灵活性大大减少。

第六，这点常被忽视，如果预先确定了雇佣的非经理人员的 APQT 束，那么对雇佣的经理人员也必须预先确定一套 APQT 束。如果雇佣的非经理人员的选择被完全控制，那末监督人员的选择也一定要被完全控制。为了保证非经理人员的 APQT 束预先确定和得到切实贯彻，经理人员的 APQT 束也必须预先确定。人们可能会问，由谁来预先确定经理人员的 APQT 束?预先确定 APQT 束不大可能普遍使用的原因，除了搜寻与监督成本高昂外，还因为它隐含着一个令全体雇员感到不舒服的组织机构。

我们可以得出这样的结论，努力是个人对内部和外部动机作出反应的结果，它受超我和本我功能的影响，也受习惯力量的影响。对习惯行为或日常行为的研究揭示出它的几个特点。第一，同样的环境刺激会导致同样——日常的或习惯的——反应，“而没有对其他各种反应的明显考虑。”第二，习惯行为会降低对环境的认识能力和反应能力，个人局限在环境提示的狭窄范围内。这里隐含着个人不大可能察觉出环境发生的变化。就是说，他们变得越来越不敏感。第三，决策

的间隔时间比较短。就是说，决策变得更加反应化①。

4.3.1 对人类认知的某些心理学研究

心理学传统可以追溯到19世纪中叶的E·H·韦伯(Weber)和G·T·费希纳(Fechner)[1860年，费希纳杜撰出心理物理学(psychophysics)一词来表示对人类认知的研究]，并经坦纳(Tanner)和斯韦茨(Swets)1954年关于"信号觉察"(signal–detection)理论的研究论文，一直延伸到较近期在这个领域作出贡献的研究——例如，埃斯蒂斯(Estes,1980年)关于决策的著作。连同其他有关的论文，这种传统导致这样的结论：人的神经系统会产生判断错误，人和其他动物的决策不但根据成本和利益，而且还根据动机和我们的（不完全可靠的）神经系统。关于这个问题，需要讨论人类感觉系统的两个特性，即差异界限(difference threshold)和适应(adaptation)。所有的感觉系统都有差异阈限。就是说，如果环境变化要被知觉的话，这个变化就必须大于某个最小量。因此，有些环境变化并未被注意到。于是，"差异阈限"(difference limen)被定义为能被注意到所必需的这种最小环境变化量。感觉系统还有适应的特性，就是说，对固定数量的能量的感觉效果会不断变化。如果你把手浸在热水里，热的感觉会逐渐被"热的中和"(thermal neutrality)所代替。正是我们感觉系统的这些"不完善性"增加了习惯行为的可能性，而减少了对量值的敏感性。当然，习惯不一定是低效率的。但是在

① 社会心理学家关于习惯和(工作)行为的论述参见韦斯(Weiss)和伊尔根(Ilgen)(1985年)的著作。

减弱我们对环境的认识方面，习惯增加了我们与环境不协调的可能性。

个人可能有偏好，但却不一定有能力或意愿对他们环境的各种变化作出明确的区分。这就是说，我们并非总是知道什么时候可以对两种不同状况作出选择。第二，虽然我们确实在按照我们的偏好进行选择，但有时候我们的习惯却会调节我们的选择和偏好。换言之，我们可能是有选择的理性，但不一定是完全理性。努力水平问题就说明了这些意思。某些努力水平显然比其他努力水平更使人偏好。例如，我们中大多数人不喜欢厌烦或过度操劳。可是，在厌烦和疲劳之间还存在许多努力水平。其中有一些很可能彼此无法区分，原因就是上面所讨论过的认知限度。因此，认为我们在难以分辨的努力水平间有所偏好是不确切的。此外，我们还要承认这样的可能性，那就是除了这种认知方面的限度外，我们并不总是按照那种作为经济人基本假设的极大化行为模式来思考问题。因此，认为我们在"难以分辨"的努力水平间有所偏爱在这个意义上也是不确切的。换句话说，不论是由于认知限度，还是仅仅由于缺乏关切来"驱动"，我们经常"并不关心应该怎样"。虽然这样的"惰性"可能在任何满足(或努力)水平上出现，但为便于说明，我们还是假设这种"惰性"只出现在相对高的满足水平上。

4.3.2 X效率理论中的惰性区域

图4-2详细说明这些思想。曲线SE表示努力水平和满足水平之间的关系。回想一下，每个努力水平至少对应一个

APQT 束或一个努力点。这条曲线可以看作有三个部分。第一部分是区域 1，也即小于等于 e_1 的努力水平。在这个区域，满足以递增的速率增加；在达到努力水平 e_1 以前，个人宁愿选择较大的努力，而不选择较小的努力。第二部分是区域 3，即大于 e_2 的努力水平。在这个区域，满足以递增的比率减小。对于超过 e_2 的努力水平来说，个人宁愿选择较小的努力，而不选择较大的努力。第三部分是区域 2，这是以努力水平 e_1 和 e_2 为边界的区域。在这个区域，努力不像区域 1 和区域 2 那样迅速地递增或递减。实际上，区域 2 显示出有一个相对的"平顶"。

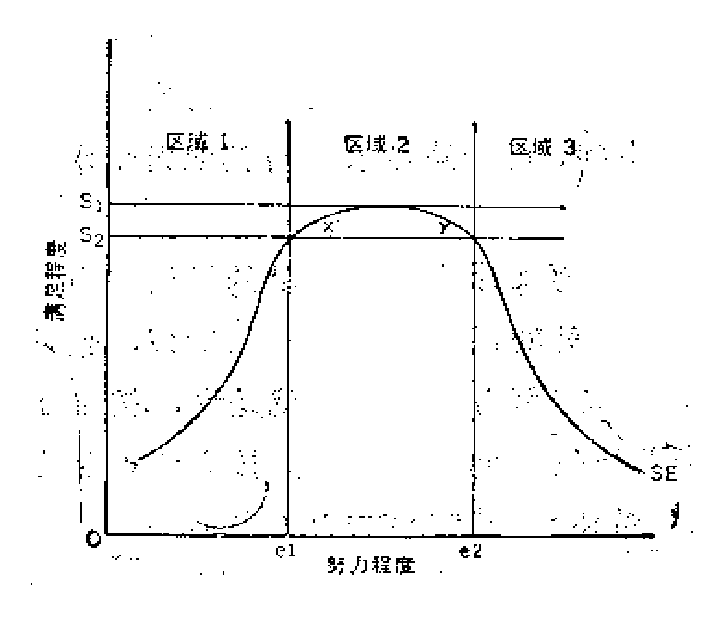


图 4-2 惰性区域

这个平顶意味着，有一个这样的努力-满足点区域，在这个区域里，每一个努力-满足点与其他每一个努力-满足点难以区分。这可能是由于上面所说的认知因素，也可能是由于个人并非极大化者这样的事实，或者是因为存在交易成本。这个难以区分的区域称为"惰性区域"(inert area)。在图 4-2 中，

惰性区域被随意画成是介于努力水平 e_1 和e_2、满足水平s_1和s_2的边界的区域。

这个惰性区域具有这样几个特性和含义。第一，个人并不总是完全理性的，即个人并不总是竭尽所能来计算和分析各种不同努力和满足水平的成本和利益的。第二，假定他(她)在计算的话，他(她)偏好的也是惰性区域内的每一努力水平，而不是惰性区域以外的任何努力水平。可见，个人宁愿要努力水平 e_1，也不愿要任何更小的努力水平。同样，他(她)也宁愿要努力水平 e_2，而不愿要任何更大的努力水平。第三，在惰性区域里的每一努力水平与在惰性区域里的其他各点难以区分。因此，努力点 X 和 Y 彼此是难以分辨的。第四，虽然人们宁愿要满足水平 s_1，而不愿要任何更低的满足水平，但是在 s_1 与 s_2 中间，每一满足水平与其他每个水平也是难以区分的。

因此个人宁愿选择惰性区域里的任何努力水平，而不愿选择惰性区域以外的任何努力水平。所以，惰性区域外的努力水平不是“均衡”努力水平，也就是说，按小于e_1的努力水平工作的个人将设法增加努力，而按大于e_2努力水平工作的个人将导致(希望)朝着较小努力方向活动。另一方面，在惰性区域 e_1 和 e_2 中间的任何努力都能是均衡努力水平。均衡努力水平的概念就有两种含义。

第一，一旦个人在他(她)的惰性区域工作，他(她)将抵制改变他(她)的努力水平。尽管任何特定的惰性区域都不一定长期存在，但在努力将超出惰性区域之前，个人仍必定会表现出“足够”程度的震惊。另一方面，由于惰性区域内的每一努力水平都能是均衡位置，所以在这个范围里，个人将是

灵活的。可见，个人惰性会产生对变化的抵制。

这里提出这样一个假设，即除罕见的情况外，惰性区域的上界——努力水平 e_3——要低于代表一项特定工作的最优理性水平的努力水平。但是，与此同时，惰性意味着趋向灵活性的一种(有限的)意愿。因此，在惰性区域内，可以指望一个个人的行动与他人的行动相吻合，不需要细致的监督就会改变努力水平并妥善地应付改变。因此，如果经理人员知道非经理人员的惰性区域，那末只要掌握他们的边界点，经理人员就能使工人达到"极大化"努力水平，而又不产生不适当的摩擦。这种摩擦包括抱怨、罢工或怠工。当然，经理们也受惰性区域的束缚。

4.4 结论

本章中我们讨论了个人的心理方面，而未涉及他(她)的外部环境。下一章将讨论外部环境的某些方面。要讨论的第一个方面是个人与他人的关系。这就要我们研究集体行为，尤其是由集体决定的努力水平。然后，我们将探索这种努力水平如何受企业在其中经营的市场结构的影响。这就是我们将要考虑的外部环境的第二个方面。我们将说明内部和外部环境同时影响X(低)效率的程度。下一章的最后部分将讨论X效率理论在生产率决定和政府管制方面的应用。

5 X效率理论 2

5.1 导言

在上一章，我们的论述集中在个人心理上，好像每个人都是鲁宾逊·克鲁索。这一章我们将研究：(1)作为集体成员的个人；(2)集体行为；(3)构成外部环境和影响企业X效率程度的一些因素。我们将从分析作为集体成员的个人开始。

5.2 个人和集体努力的企业内部的决定因素

个人的理性程度可能是他（她）自己独立"选择"的结果，也可能是受他人影响的结果。这些"有影响力的他人"①可能是他的同事，也可能是他的上司。莱宾斯坦把这些关系分别称作为"水平的"关系和"垂直的"关系。第三种可能性是，个人还会受到过去发生的，现在已经成为企业传统和历史的事情的影响。不论是否正式，传统都能够而且常常会对

① 消费者行为理论中所谓的同事集体压力，在早川(Hayakawa)和韦尼耶里斯(Venieris)(1973年)的著作中有论述。

行为产生十分有利的影响。传统对行为的影响经常采取下面的这种问答形式。问："你为什么这样工作？"答："我不知道，我们一直就是这样做的。"也许"这样做"就是最佳的，但是，并没有为什么应该这样做的先验理由。①

5.2.1 传统、上司和同事

这些其他影响的效应如图 5-1 所示。图中用曲线 SE 表示一个特定个人的满足-努力关系，用 e_1e_3 表示他的惰性区域。回忆一下上一章所述，惰性区域表示不存在充分理性的决策。因此，惰性代表这样一种状况，在这种状况下，行为具有习惯、不注重细节和对环境改变相对缺乏认识的特征。再回忆一下，惰性区域可以在任何努力水平和满足水平上出现。但是，为了便于说明，我们假定它出现在相对高的满足水平上。

以满足-努力关系 SE 和惰性区域 e_1e_3 为初始状况，现在，我们根据垂直关系对关系 SE 的影响来调整这个关系。在这样做时，为了便于说明，我们提出两个假设。第一，我们假定垂直关系会对较大努力造成压力，而个人则可以通过避免这种压力得到满足。第二，我们假定个人也可以从上司的称赞中获得满足。换言之，大多数人都能因避免"过分拚搏"和所做的工作获得别人赞赏而得到某种满足。

如果个人在"低"努力水平工作时能从避免来自上司的压力中得到一些满足的话，那末在低努力水平上，"经过调整的"

---

① 关于各种非极大化的决策技术，可参见莱宾斯坦(1983年，1985年)。

满足-努力曲线SE将处在“未经调整的”满足-努力曲线SE的下方。如果他从对他高努力水平的称赞中也获得一些满足的话，那末SE′就将处于SE的上方。因此，就相对低的努力水平来说，SE′低于SE是因为来自上司的压力会使满足减少。同时，对于相对高的努力水平来说，SE′高于SE是因为上司的称赞使满足增加。因此，上司的压力促使个人更愿意努力地工作，即变得更加理性。①

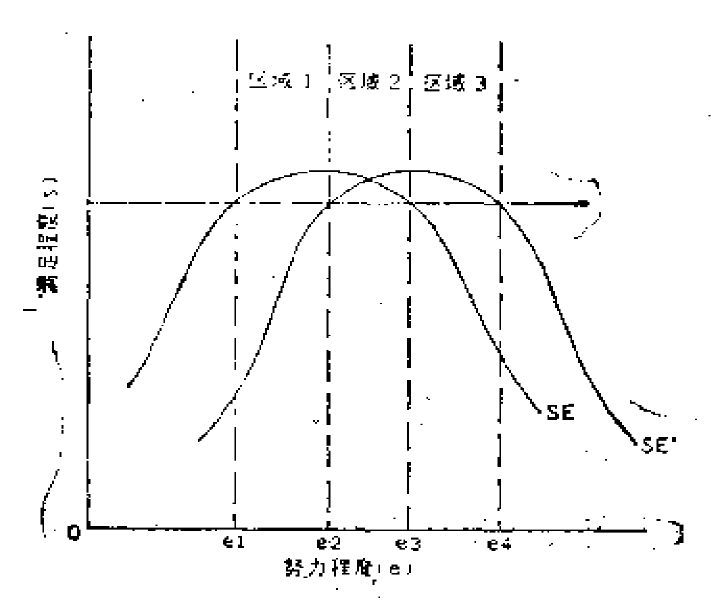


图5-1　垂直关系和惰性区域

经过调整的关系SE′形成它本身的惰性区域e_2e_4。正如预期的那样，对于未经调整的惰性区域e_1e_3，经过调整的惰性区域的下限(e_2)和上限(e_4)都是处在较高的努力水平上。因此，在这个上司-下属关系中的个人惰性区域是e_2e_4。很显然，图5-1的区域1——努力区域e_1e_2——将不再是惰性区域的一部分。知道自己在区域1工作的个人将发觉在这里工作的

① 较高和较低努力水平的相对压力不必彼此成比例。因此，SE和SE′的斜率和宽度不一定彼此相等。

不舒服足以促使他进入区域2。区域2肯定在新惰性区域内，因为它是SE和SE′“共有的”。因为受称赞增加满足，所以预期区域3也在他(她)的新惰性区域内。但是在这里，还可以讨论一个新的因素，对于人的努力水平来说，这个因素的含义饶有趣味。

莱宾斯坦指出，在区域3中SE′位于SE之上，因为个人从称赞中会得到满足。但是在区域3，SE′上的各点还应该予以考虑，因为个人并不是他(她)自己的“主人”。这就是说，他(她)在区域3工作，仅仅是为了得到他人的称赞——在这里，仅仅是为了得到上司的称赞。因此，在区域3工作既有被称赞的满足，但也有不能自己作主的不满足。如果这两股力量的净效应为正，那末区域3就在新的惰性区域之内。因此，就像画出的那样，图5-1是假定净效应为正。另一方面，如果净效应为负，那末，经过调整的SE关系就显示为图5-2

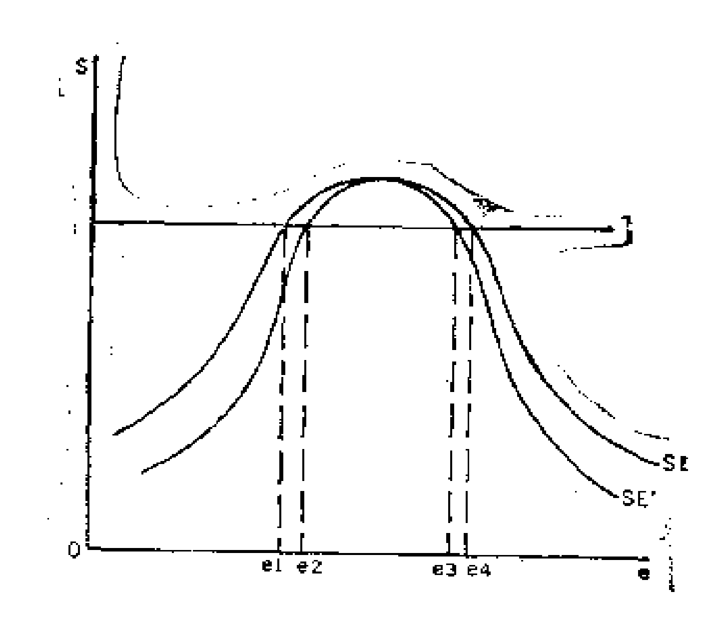


图5-2　存在由非自我雇佣产生的净不满足时的垂直关系和惰性区域

中的SE′。在这种情况下，新惰性区域是$(e_2e_3)+(e_3e_4)$或e_2e_4。这种情形的一个含义是，经理们只要创造一个使雇员们感到他们是在为他们自己工作的环境，并使雇员感到他们在企业生活中是与他人平等的合伙人，这样就能提高惰性区域的上限。

在可以预期上司试图会增加努力的时候，同事的压力将产生什么效应还是不清楚的。我们描述三种可能性。第一种是"下限"压力，那就是集体的全部成员预期都会"尽到他们的职责"，而且在他们这样做的时候，他们可以不受限制地想尽多大力就尽多大力。第二种是"上限"压力，也就是个人只要愿意就可以少做，但他们却决不可以太努力，以免使别人显得不努力。第三种是上下限压力，那就是你尽你自己的职责，但不要工作太努力，以免使他人显得不努力。

在图5-3中，SESE′指的是未经调整的惰性区域，e_1e_4就是其惰性区域。因为压力只落在下限上，所以可以预期，个

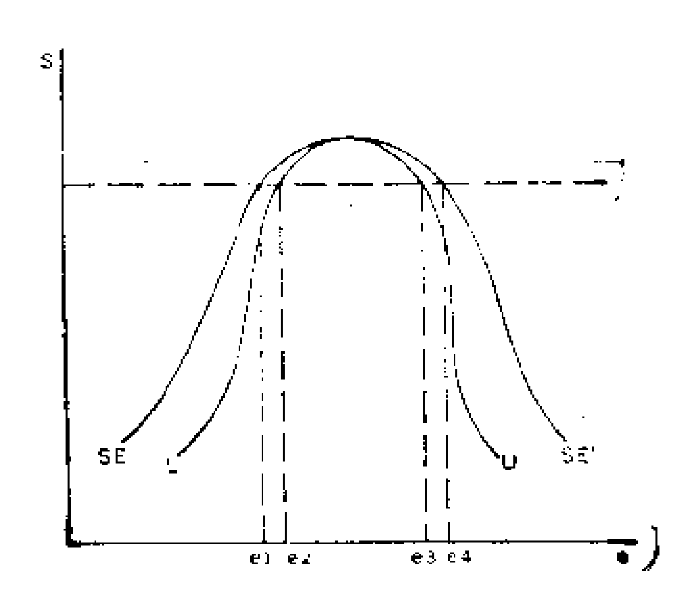


图5-3　工作习惯和惰性区域

人只在这些较低的努力水平上才会感觉到压力。当我们调整SE曲线以考虑同事压力的时候，经过调整的关系如LSE′所示，新的惰性区域成为e_2e_4。

上限压力也如图5-3所示。这里未经调整的关系为SESE′，惰性区域为e_1e_4。经过调整的关系使关系SESE′变窄，成为SEU，使惰性区域缩小为e_1e_3。最后，在图5-3中，上下限压力并存的情况表现为LU，惰性区域为e_2e_3。

5.2.2 多人努力和惰性区域

多人惰性区域的概念是从个人未经调整和经过调整的惰性区域概念得出的。多人惰性区域由每个集体成员共同惰性区域内的全部努力点构成。在图5-4中，我们假定这个集体由两个个人组成。个人A经过调整的SE关系以SEA表示，个人A的惰性区域为e_1e_3。个人B经过调整的SE关系和惰性区域分别为SEB和e_2e_4。因而，集体的惰性区域为e_2e_3。也就是说，在e_2e_3区域内，两个人都没有施加压力去改变他自己或另一个人的努力点的意向。个人A不欢迎个人B表现出任何高于e_3的努力水平，并会作出使B减少努力的反应。同时，个人A表现出任何少于e_2的努力水平都会引起B作出使A增加努力的反应。然而，只要A至少按努力水平e_2工作，而B按最大努力水平e_3工作，那末A和B就都不会有改变他自己或对方努力的动机。这样说并不意味着A或B都处在他(她)最偏好的努力水平上：任何一方都可能对这种现状感到不"快乐"，但他们却谁都不想去改变它。

随着集体的规模增大，集体努力将具有公共货物的特性。

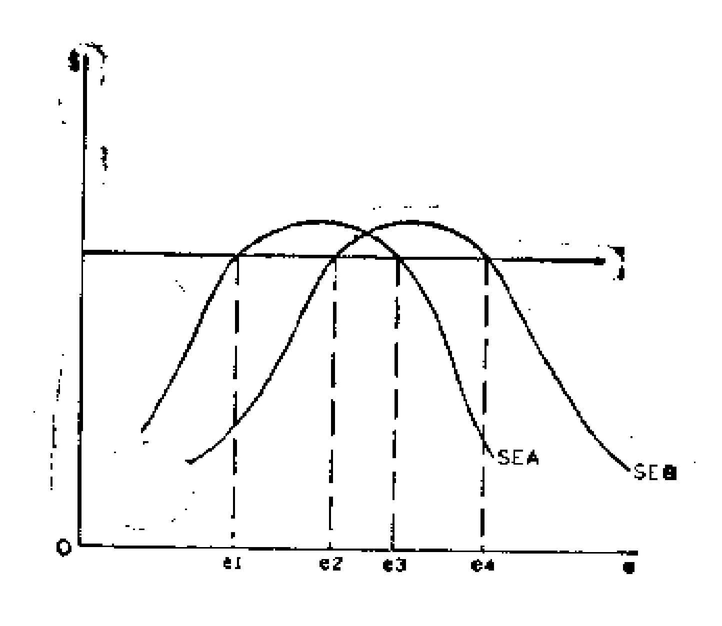


图5-4　多人惰性区域

如果其他条件不变，我们可以预期，集体成员人数的增加将缩小集体的惰性区域范围；也就是说，集体成员人数的增加将减少每个人都可以接受的努力点的数目。

另一方面，不难想象，集体成员人数的增加将导致比较“正规的”集体工作道德规范，诸如上限或者下限压力，或者两者兼有。在这种情况下，每个成员就都会同意“相互容忍”，只要其他个人不破坏这个规则。结果是惰性区域范围的扩大。在图5-4中，变得更有“满足弹性”的每个集体成员的SE关系都显示出这种情况①。这里不准备讨论关于集体规模和惰性区域的准确后果。只要说清楚集体关系对努力程度会产生影响就足够了。

多人惰性区域的概念也意味着相对较低和较高的努力水平将不再是集体惰性区域的一部分。因此，集体表现出的努

① 集体内聚力和生产率在弗朗茨和格林(Frantz and Green，1982年)著作中有论述。

力水平将接近于每个集体成员的"平均"生产率水平。因而，我们可以认为，集体努力是一种"协定"(convention)或协议(agreement)。下面我们探索它对生产率的决定有何影响。

5.3 生产率、努力协定和囚徒的两难困境

个人努力的相机抉择(discretion)概念和多人惰性区域概念意味着企业中没有一个人能控制决定努力程度的所有相关变量。在这种情况下，努力或(生产率)可以看作是一局对策的结果。莱宾斯坦使用囚徒的两难困境对策来说明生产率的决定问题①。简要地说，囚徒的两难困境是这样一种情况，即每个个人照他(她)的自我利益采取行动，结果对任何个人都不是最优的。换一种方式说，囚徒的两难困境在于，对手的行为机会使合作的(优超)行为减少。

表 5-1 表示作为囚徒两难困境对策的生产率。该表表明经理人员(企业)和雇员关于对方各有三种选择(option)。每方都能采取这样的一种策略，那就是：(1) 黄金法则(GR)；(2)同业集体标准(PGS)；或者(3)个人极大化(IM)。对于企业来说，黄金法则意味着企业经营几乎完全是为了雇员的利益。黄金法则策略是一种合作策略。另一个极端是个人极大化，或成本极小化。在这一端，企业试图使提供给雇员的工作条件(包括工资)极小化，而同时他们又试图使能从雇员那里得到的东西达到极大。这个策略是完全敌对的策略。第三种选择或"中间道路"是，提供的工作条件符合"做一天公平工作，付一

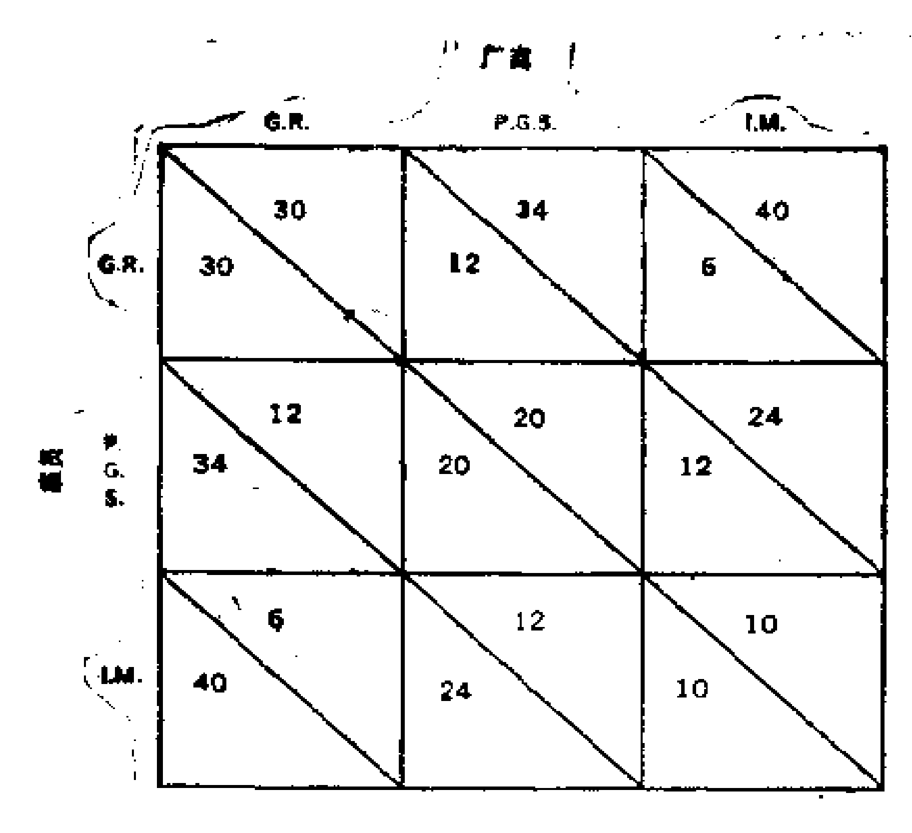
① 这点在莱宾斯坦(1976年，1982年)著作中有论述。

天公平工资”的哲学。正如我们将熟悉的，这也是一种合作策略。

对雇员来说，黄金法则策略就是他们像企业的主人一样工作。也就是说，他们根据企业的最大利益工作，他们希望企业如何对待他们，他们就以同样的态度对待工作。雇员的黄金法则策略也是一种合作策略。另一个极端是，尽可能小地“付出”努力，而工资照拿。这个策略是完全敌对的。从个人说，可以把它看作是“免费搭车者”的选择。就是说，个人的逻辑是，只要他(她)是唯一的“偷懒者”，企业就不会因此遭受损失，因而偷懒也就不会引起注意。免费搭车对任何个人来说都是一种理性的策略，因为它使工作努力极小化，却又没有利益(工作条件)的损失。中间道路就是“付一天公平工资，做一天公平工作”的哲学。再说一遍，可以证明这是一种合作策略。

这局对策的囚徒两难困境方面在于，如果每一方都遵循自己的极大化策略，那末双方都会有“损失”。就企业而言，成本极小化要求不管雇员使用三个策略中的哪一个，企业都应该尽可能地少给。就是说，企业应该永远采取与个人极大化相敌对的策略。成本极小化(利润极大化)要求企业对于雇员既定的努力水平尽可能地少支付。与此同时，对于企业采取的任何既定策略，雇员的理性策略都是尽可能少地付出努力。在这两种情况下，使自己的利润/成本比率极大化，都要求采取个人极大化的敌对策略。结果，当然是囚徒的两难困境：一个低生产率和工作条件可怜的企业。

幸运的是，我们中的大多数人并不是理性的！表5-1的数字代表每一方从每次联合选择获得满足的值。因此，如果



	厂商 G.R.	P.G.S.	I.M.
G.R.	30 / 30	12 / 34	6 / 40
P.G.S.	34 / 12	20 / 20	12 / 24
I.M.	40 / 6	24 / 12	10 / 10

表 5-1　作为囚徒两难困境的生产率

双方都选择黄金法则策略，那末企业的满足（获得一定水平利润的满足）和雇员的满足(获得一定工作条件的满足)都等于30。另一方面，囚徒两难困境解的每一方的值都只有10。企业的最好联合选择在右上方的方格里。也就是说，企业的最好选择在于，在企业对雇员的支付达到极小化的时候，雇员能按照黄金法则工作。换言之，企业的最佳地位是，当它采取敌对行为时，工人们却抱合作态度。这些工人或者只有很少的讨价还价能力，或者是一批“努力工作的傻瓜”。这个选择对企业值满足水平40，但对雇员只值满足水平 6 。

另一方面，雇员的最佳联合选择在左下方的方格里。就是说，雇员的最佳选择是，在他们受到第一流“待遇”的时候，却以免费搭车行为作为回答。换言之，雇员的最佳联合选择

是，当他们采取敌对姿态行动时，企业却采取合作策略。如果其他条件不变，在这个联合选择中，企业的利润将是十分低的(高成本加上低生产率)。这个选择产生的满足水平对雇员为40，而对企业却只有6。(虽然表5-1中的支付是对称的，但在联合选择中这样的对称却不是必然的。)对角线以外的其他方格也代表其他敌对策略的支付。

值得注意的合作联合选择是同业集体标准。在这个联合选择中，双方的满足水平都是20。莱宾斯坦把同业集体标准的解看作是囚徒两难困境解的一种替代。如表1-1支付矩阵所示，同业集体标准虽然不如黄金法则解那么好，但却优于囚徒两难困境解。如上所述，同业集体标准代表一种努力习惯(convention)或规范(norm)。关于努力习惯有两个问题：(1)它存在的理由；(2)它如何影响行为。努力习惯的存在理由是戴维·刘易斯(David Lewis, 1968年)和埃德娜·厄尔曼-玛格丽特(Edna Ullman-Margalit, 1977年)阐发的。一个主要的理由是习惯可以解决协作问题。

努力习惯的特别引人注意之处在于，不止一种习惯能解决企业的协作问题。在这种情况下，任何习惯不管是最优的还是非最优的，都比没有习惯的强。任何习惯好处的一个说明(莱宾斯坦曾引用过)是，所有汽车该沿路的哪一侧行驶的决策。

一旦努力习惯形成，大多数人就会遵守它。新成员观察更有经验的老成员的APQT束，把它作为学习"如何做事"的一种方法。如果有人感到努力习惯不舒适，那么很明显，他(她)就只好或者调整自己的标准，或者另谋他就。但是一旦接受，个人对这种习惯的反应就具有"刺激-反应"的形式。那

就是，每当要求运用习惯时，不必对它再作评价。习惯是为运用而存在的。它能加以重新估价和变动吗？当然可以，但只有在完全意想不到的冲击使个人震惊得离开他们的惰性区域时，这种情况才会发生。

现在回到表5-1，显然经由同业集体标准的努力习惯产生的合作策略，虽然优于导致囚徒两难困境解的敌对策略，但却不如黄金法则的合作策略。因此这种努力习惯还不是“最优”解。那末为什么个人和(或)集体都同意遵守这种并非最优的习惯呢？一个理由是，实施得很好的制裁办法将加诸任何不遵守习惯的人。集体中任何个人表现出的努力水平与其说是接近于他的最好水平，不如说更接近于平均的努力水平。同业集体标准解的另一个特点是，除了黄金法则解，它既“有效”又“公平”。囚徒两难困境解显然没有效率。另一方面，不在对角线上的要素至少在一方看来是“不公平的”。例如，雇员会认为，他们按照黄金法则工作，而企业却表现为个人行为极大化的情形是不公平的。同时，企业会认为它按照黄金法则，而雇员表现为个人极大化行为的情形是不公平的。其他不在对角线上的要素也代表一方的付出超过另一方的情形。另一方面，同业集体标准解是一种双方付出和收入“相等”的解。

5.3.1 福特汽车公司，路易斯维尔工厂：离题话①

描写工商业的作品中有许多关于企业“转变方针”(turn

---

① 梅因(Main，1983年)讨论了这个例子。

round)的例子。路易斯维尔的福特工厂，就是一种比较互相有利的解怎样取代囚徒两难困境解的一个例证。至少在1979年以前几年，福特公司在路易斯维尔的装配厂就以"战区"闻名，那里厂地污秽，劳工与管理人员关系公开敌视，经理人员专制，工会领导人毫不让步，互不信任和极不关心X效率。因而未到1979年，该厂汽车需求锐减，夜班停产。谣言纷传工厂即将关闭。

就在这时，该厂工会主席和经理易人，新上任的两位相信，自毁行为是没有必要的。事情很清楚，如果不想关门，工厂就必须在六个月内"转变方针"。少数重要的雇员要求在工厂午餐。公司实施"工作生活质量"计划，不断与雇员进行认真的讨论，搞了许多虽小但很有意义的改革。一项大的变革是取消设计者把设计送给将要按其设计进行生产的产品制造者的传统做法，这常常由于制造者试图遵照(不必要的)困难的设计而不必要地付出大量努力。福特的路易斯维尔工厂打破传统的做法是，把那些设计在厂内公布，在制造阶段开始之前就征求工人的意见。到1980年初，这家工厂生产出高质量的产品，到1980年底，福特公司决定在福特兰杰牌汽车的生产上投资7亿美元。1979年以前的年月里，每100辆汽车中质量检查员记录下的"质量问题"多达700个，到1982年，这个数字下降为198，成为福特公司在美国历史上的最低记录。再次，很明显，会失去工作引起的"冲击"，使囚徒两难困境解转变为双方更能获利的某种解。工人和经理人员在1979年前与在1979年以后表现出的行为，是在追求他们的效用极大化吗？虽然这可以是他们行为的一种解释，而且它也符合新古典理论，但我却认为这种解释是很牵强的。这个问

题和类似问题将在第9章讨论。

5.4 市场结构、压力和努力

5.4.1 导言

在X效率理论中，企业购买的是工人的时间，但使用的是他们有目的的努力，工人有努力的自由处置权，可以任意选择他们的努力水平。这些个人的努力水平受到这样一些惰性区域的约束，这些区域考虑到了他们自己的选择会如何影响别人，以及如何可能受别人的影响。换句话说，集体惰性区域是每个人都"赞同的"努力区域。每个雇员带给企业的不只是时间和努力，而且还有一套影响企业内"气氛"的特征。这种气氛的一个要素就是激励机制，激励机制是一种企业通过市场机制无法买到的"投入"。这种机制决定用什么刺激企业内个人与集体更加努力，更加理性。这种激励机制不但取决于企业内部各种因素，而且取决于企业外部各种因素。在这一节里，我们要考察这些因素。

5.4.2 垄断企业

垄断力量使企业成员不再受到改变他们均衡努力位置的压力。可以证明，垄断企业对于价格、产出和成本的控制，是造成这种状况的原因。为了说明可以通过相对高的成本、价格以及较低的产出获得既定数量的利润，人们可以编出一系列关于价格、产出和单位成本的数字。换言之，一旦我们放

弃利润极大化——成本极小化——行为的假设，我们就能得出这样的结论，即如果可接受的利润不必达到成本极小化就能获得，那么成本极小化就不是垄断企业中的典型行为。如果雇员有任意决定努力水平的自由，那么垄断力量就会使雇员与任何降低成本的压力隔绝。只要企业满意它的利润约束，那么就没有理由相信（先验地）成本极小化是垄断企业的典型行为。（在下一章中，我们将考察包括利润约束的几个模型。）这里的假设是，如果企业成员能把他们的惰性"传递"给（广泛分布的）所有人和消费者的话，那么他们就一定会这样做。在垄断条件下，这一点肯定是可能的。

资本市场：简短的离题话

经常可以听到这样一种反对论调，认为即使产品市场被垄断，完全的资本市场也将会迫使雇员们（经理们）追求成本极小化，否则他们就会被接管，由成本意识较强的经理人员取而代之。因此，可以认为，完全资本市场的存在已使得X低效率不可能存在①。奥尔钦和凯塞尔（Alchian and Kessel，1962年），迪·洛伦佐（Delornzo，1985年），法马（Fama，1980年），曼纳（Manne，1965年）都以不同的方式表述了这个观点，有的还直接提到了X效率理论。资本市场完全吗？证据是很不充足的。例如，罗伯特·斯迈利（Robert Smiley，1976年）对1956—1970年间美国95个股权收购案例的研究报告认为，公司帐面价值超过市场价值约13％之前接管是不大可能发生的。奥利弗·威廉森（Oliver Williamson，1970年）报道说，这个"低效率界限"

---

① 关于资本市场和X效率之间的关系，见弗朗茨（Frantz，1984年）的论述。

在10%-25%之间。不管为什么会存在这样一个低效率界限，对于企业成员来说，它的存在减轻了他们提高X效率，保持警觉和成本意识，以及重新估价它们均衡努力位置的压力。莱宾斯坦使用的“努力熵”①这个术语没有受到非难。

除了存在一个低效率界限外，辛格(Singh, 1975年)和弗思(Furth, 1980年)使用美国与英国的资料进行的研究报告说，不但无法保证兼并后的较高利润，而且在美国，被兼并企业的经理和董事常常还会得到很大的利益。合在一起，这两个“事实”提出了一种饶有趣味的可能性：X低效率不但可以泰然容忍，而且在某些条件下它还可能是“理性的”。

企业家才能：又一番离题话②

X低效率不可能毫无改变地维持下去。成本增加和利润减少导致压力增大，而且新的竞争带来的威胁，对工作和收入造成了一种严重威胁。在减少X低效率方面，企业家的作用表现在两个方面，一是对低效率企业的直接影响，二是作为一个竞争者的间接影响。

另一方面，正如第2章提到的那样，企业家才能是循规蹈矩和无关紧要的。如果市场是完全的，如果企业像例行公事那样以极小成本把它的投入转变为技术上的极大产出，那么企业家才能就只是按惯例执行的职能，并不重要。新古典理论中的成本极小化假设，也许就是企业家才能之所以相对不受重视的一个原因。但是在X效率理论中企业家才能是重要的，因为这个理论假定市场不完全，企业不会实现成本极

① “努力熵”一词，莱宾斯坦(1976年)曾有论述。

② 关于企业家才能，莱宾斯坦(1978年)曾有论述。

小化。

不完全市场一词在这里有两个意思。第一，安排(marshalling)和利用投入存在一些障碍。莱宾斯坦提出了这样一些例子。企业可能获得的某些投入现在却正忙于其他工作。为了找出有关这些投入的另外可供选择的结构和安排，这就需要想象，而不单是计算。第二，投入可能存在，但却不是按已知的特定价格销售。第三，投入的价值只有考虑到能生产以前未受重视的产品的整组投入时才能确定。第四，某些投入，如资金，有些人能够得到，但另一些人可能得不到。

市场不完全的第二个意思是，投入和产出市场上存在各种不同类型的"漏洞"。例如，知识之类的投入并不总是可以进行交易的。把知识从感觉或想法变成对企业有用的形式需要激励。因为，生产函数被假定是不完全的，因此，可以把企业看作是受到适当激励促动就会发生作用的经验仓库。在欠发达国家，可以把资金加到并不总是进行交易的投入清单中去。

在这种情况下，莱宾斯坦将经济看作是路径与结点组成的"网"。结点是那些得到投入的企业和行业(以及得到消费品的那些家庭)。路径则把投入送到企业，再把其产出送往其他企业(和家庭)。在遍布成本极小化企业的功能良好的市场中，这张网是完整的，没有任何漏洞，路径也很明确；而且对于相同的商品，每个结点都在平等的基础上与其他结点进行交易。在更符合实际的情况下，企业家的日常职责就是充当一个改进路径和结点的"投入完成者"和"缺口填补者"。此外，企业家在发明与创新中的作用可以看作是对这张网的扩充。莱宾斯坦的结论是，企业家在市场间有许多方式可以发挥作用。

因此，X低效率的存在有助于激发企业内部和作为竞争者的(潜在的)企业家才能。根据上面的结论，企业家才能的提供不一定假设是完全有弹性的。

5.4.3 竞争企业

其他情况不变，竞争对比较理性的企业成员会产生更大的压力。这个效应是通过市场或价格机制产生的。依靠这种机制，如果一家企业降低它的价格，那么所有企业就必须跟着降低，否则就会被这个行业淘汰。在长期，企业靠尽可能压低成本得以生存，它不但要利用所有的规模经济，而且还要能在长期平均成本函数(曲线)上进行生产。因此，竞争减少了(某些)垄断企业雇员采取各种形式的自由处置行为——包括专断的、草率的、官僚主义的、傲慢的和不作反应方式的行为——的机会。在竞争条件下，对其他雇员和顾客的高度责任感，是保证企业以极小成本为市场服务的一个特征。企业将更加彻底地检查和更加经常地重新估价它的努力点，因为生存的压力迫使它做这样的重新估价。

5.5 综合和实例

5.5.1 耶基斯—多森法则

人格与环境就这样"组合起来"产生X低效率。不论是在人格还是在环境情况下，压力都是关键概念。因为存在超我和本我需求，人格创造出它自己的压力。环境通过企业内部

人际关系和市场竞争(或缺乏市场竞争)的效应产生压力。为了说明压力与X低效率之间的关系，莱宾斯坦改写了1908年实验心理学家提出的耶基斯—多森法则。这个首先由罗伯特·耶基斯和约翰·多森(1908年)写下的"法则"(通过老鼠实验)揭示了刺激强度与学习之间的关系。在某些情况下，只有在一定的刺激水平上，学习才随着较大的刺激增加；而随着刺激的继续增加，学习将逐渐减少。这就是说，这种关系存在两次关系的特征。以后，这种关系被解释为动机(或刺激)与工作绩效之间的关系①。莱宾斯坦自己对这个法则的改写，是要讨论压力与工作绩效间的可能关系。他的假设是，这种关系可以二次方程形式描述。即承受相对较低或相对较高水平压力的个人，是不会尽可能做好工作的。另一方面，存在一种"中间的"(最优)压力水平，在这种压力下，个人会有最好的工作绩效。

图5-5中，莱宾斯坦改写过的耶基斯—多森法则表现为压力与决策质量之间的平方关系。图5-5表明，压力水平p*导致最大绩效。比p*小的压力水平代表一种"安逸"的环境，它不会提供使个人利用他(她)的技能所必要的刺激。另一方面，比p*大的压力代表水平太高的、不能导致最优绩效的压力。因此，只有压力水平p*才会产生最优或极大化行为。因此，压力水平p*(如果其他条件不变)将使X低效率达到极小。因此，这里极大化一词是在非极大化有可能出现的意义上使用。非极大化不一定被观察到。但是，它被认为是人类行为的可能结果。

---

① 莱宾斯坦(1983年，1985年)讨论了耶基斯—多森法则。西托夫斯基(1976年)也讨论了刺激与行为之间的关系。

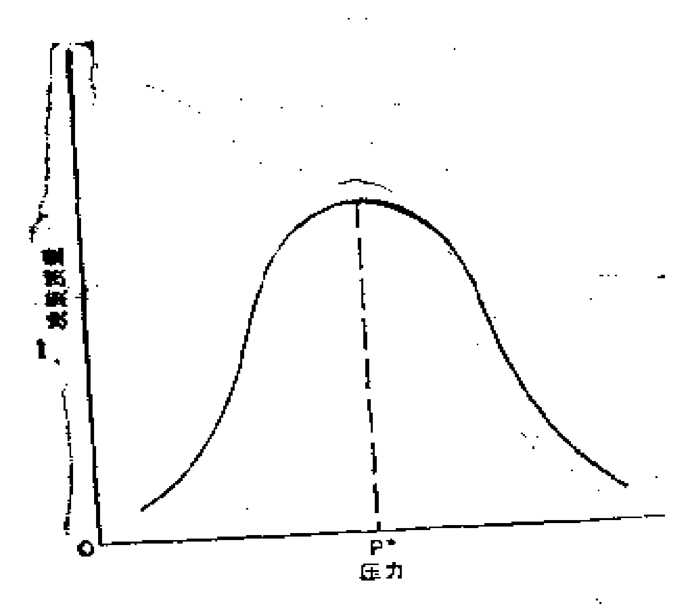


图 5-5　改写的耶基斯--多森法则压力与决策质量

5.5.2　综合

图 5-6 试图说明本章与前一章提出的许多观念。在象限1,压力是自变量，而个人的关心约束水平是应变量。这种关系有两个特性。第一，不管压力强度如何，个人能够持有的最高关心约束水平即现实主义总是存在的。第二，为了描述惰性区域，这种关系不用线条而用“宽”带画出。在这种情况下，对任何压力水平，个人能够持有的关心约束水平都有一个范围，范围内的每一个关心约束水平都与任何其他水平一样令人满意。在象限 2 中，关心约束水平是影响努力的自变量。换言之,被称为关心约束的决策特征会影响工作行为。除了说明惰性区域，任何关心约束水平都与一定范围的努力水平相一致的关系表明，随着关心约束水平增加，努力的供给将变得更有弹性。换句话说,就是存在一个关心约束水平,在

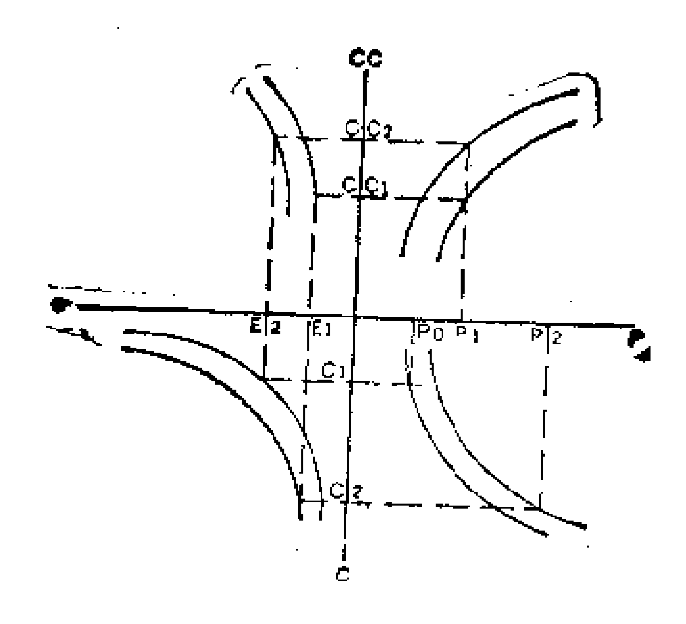


图 5-6　综合：X低效率的决定因素

这个水平上努力足以使X效率达到极大。

在象限 3 中，单位成本表示为努力的函数。象限 3 表明，不管努力水平如何，都存在一个“固定的”成本水平。它还表明当努力趋近于零时，成本将趋向无穷大。象限 3 说明，努力水平对成本具有非决定性的影响。在这种情况下，我们就能够理解，象限 3 所示的每一个努力水平都会导致不同的成本函数。最后，象限 4 表明，压力是成本的函数。也就是说，成本上升会增加企业(成员)减少成本的压力。而且，成本增加时，压力还会以递增的速度增加。垄断企业比竞争企业更缺乏压力的成本弹性。总之，对于任何成本水平，个人都可以察觉到一个压力范围。这就是说，这个关系可以看作是包含了一个惰性区域。

因此举例来说，这个关系的顺序是这样的：一定的压力水平导致某种关心约束水平，关心约束水平影响努力，努力影响成本，成本转过来又影响压力。在图 5–6 中，我们任意地

从压力水平 P_1 开始。P_1 导致关心约束水平在 CC_1 和 CC_2 之间。接着，它造成的努力水平在 E_1 和 E_2 之间。这个范围的努力导致成本处于 C_1 和 C_2 之间，该成本水平转过来又造成个人能觉察到的压力处于 P_0 和 P_2 之间。如图 5-6 所示，在一系列同时存在的关系中，惰性区域的概念意味着应变量的值是"模糊的"，包含着一定的不确定性。压力与成本的最终均衡水平取决于每条曲线的位置和弹性。我们所说的只是许多可能的动态结构中的一个。

5.5.3 实例：克莱斯勒公司的"底特律汽车装潢"工厂

在 1983 年 4 月 4 日出版的《幸福》杂志中①，为克莱斯勒公司制造座套的"底特律汽车装潢"工厂被说成是一家"十分普通"的工厂。底特律汽车装潢厂的问题，和克莱斯勒公司在底特律的其他工厂一样，到 70 年代末和 80 年代初就变得很明显了。1981年，克莱斯勒公司销售额为 100 亿美元，亏损 4.76亿美元。同年，该公司改变预算制度，对像底特律装潢厂那样的工厂的雇员形成很大的压力。新制度不再让每个工厂根据它的成本制订预算，而是要求他们将自己的成本同克莱斯勒公司从别处购买同样数量与质量的产品所花的成本进行比较。底特律汽车装潢厂 1983 年度的成本达 5150 万美元，而同样的座套向外面的卖主购买只需 3080 万美元（相当于打 40% 的折扣)。原来估计为克莱斯勒公司赢利 100 万美元的工厂，不可能以每年损耗 2000 万的代价维持下去。实质

① 梅区(1983年)论述了这个实例。

上，底特律汽车装潢厂是在接受公司不再有能力承受的补贴。

于是克莱斯勒公司请阿瑟·D·利特尔公司(Arthur D. Little)复核，其费用85000美元由工人和经理人员平均负担。利特尔公司证实了3080万美元这个数字，工厂雇员的整个态度立即发生变化。很明显，工厂必须有竞争能力，否则就只好关门。在劳资谈判中，工人们清楚地看到，工厂冗员太多，工人、经理部门和工会人员都有过剩；不必要的工作规则也成为负担；雇员经常每天只做三四个小时的工作。为了提高工厂的竞争力，建立了一个由工人和经理人员组成的14人整顿委员会。克莱斯勒公司要求减少640万美元成本才能保全工厂。而整顿委员会的计划是节省540万美元，克莱斯勒公司同意了。该计划要求裁减全厂人员的25%（从709人减为528人)。"直接"(生产)人员、薪金人员、"间接"(杂务等)人员分别裁减21%、25%和40%。工头与工人的比例从1:21增加到1:30。公司付薪的工厂工会人员由6人减为4人。工作班组人数减少，工作规则更加自由。每天的产出指标视工种不同分别提高15%至28%不等。产量增加只作为保住工作的手段而受到"欢迎"。另一方面，在座垫背上缝滚条的工人将要求每天工作量从293只增加到376只——增加28%——在此文写作时，一个女工的产量已达到330只。开始时，提高努力是困难的，但是，提高的努力水平的保持时间越长，困难似乎就越小。到该计划的80%付诸实施时，每年节约的数字已达到了640万美元。据利特尔公司的代表说，美国许多工厂都有与底特律汽车装潢厂同样多的改进余地。底特律汽车装潢厂生存下来后，它的经理总是提醒人们，"不要

忘记那个松松垮垮几乎关门的工厂。"①用表 5-1 解释，底特律汽车装潢厂的变化说明，在施加了足够压力的情况下，囚徒的两难困境能转变为较好的解法，虽然我们的资料还不能决定是哪一种解。

5.6 含义和结论

X效率理论的几个含义是：(1)企业的生产函数是不完全知道的；(2)企业不仅仅是一个生产函数(产出根据即定投入变化)；(3)成本不一定极小化；(4)成本是价格(市场力量)的一个函数；(5)人的努力不一定加以完全利用，即如果运用适当的刺激，收入和工作的满足都能够增加。

X效率理论的提出是为了解释看来与当时正统微观经济理论不相一致的实际资料。与其他经济学家如鲍莫尔、马里斯和威廉森的研究方法不同，莱宾斯坦试图通过研究一般性的问题来解释企业行为。这些问题有：是什么决定工作努力；在什么条件下个人会向着X效率努力，即会尽力于充分利用企业所支配的资源。这种方法的一个结果是产生个人不一定是程序意义上的极大化者这样一个概念。这一点比任何其他方面都更使X效率理论与新古典理论有所区别，并引起了对它的强烈反应。这种反应将是第9章讨论的主题。

在随后的三章中，我将提出一些经验研究以支持X效率理论。读完这三章的读者可能会发觉这些章节不但有助于他(她)注意到这些研究怎样与新古典理论相一致，而且也有助

① 见梅因(1983年)的论文。

于他（她）注意到何种类型的研究设计能够区分这两种互相对立的解释。

附录：管制与X效率

A.5.1 导言

许多作家认识到X（低）效率的（潜在）重要性，但没有提出任何经验性证据。例如，鲍莫尔和克莱沃里克（Baumol and Klevorick，1970年）说，尽管主要兴趣在报酬率管制的配置效应上，但X低效率效应都很可能是大量低效率的根源。克罗斯（Cross，1970年）关于公用事业"激励定价"的论文也承认X效率的重要，尽管它被经济学家普遍忽视。对受管制企业中存在的X低效率和配置低效率，约翰逊（Johnson，1970年），马斯金和索伦蒂诺（Muskin and Sorrentino，1977年），谢尔曼和威尔逊（Sherman and Wilson，1976年）都有论述。怀尔德和斯坦塞尔（Wilder and Stansell，1974年），也引证了公用电力事业中X低效率的重要性，而巴伦（Baron，1974年），梅纳德（Maynard，1979年），纽豪斯和泰勒（Newhouse and Taylor，1970年）都写过医疗保健中的X低效率。

鲍莫尔和克莱沃里克提出了两个说明管制为什么可能增加X低效率的理由。第一，人们相信，管制者会利用报酬率管制使低效率企业留在行业内。这就是说，管制者允许低效率企业通过以较高的价格这种形式把（至少部分）低效率转嫁给顾客而取得利润。第二，管制者对利润设置最高限额，往往严重地削弱了创新和效率的激励，换言之，削弱了对"杰出

企业家成就"的激励。

A.5.2 价格和效用，或利用不足

因此一般说来，政府管制工商业，保护企业(及其雇员)免遭竞争，会降低对直接努力的激励，并招致X低效率。这就是说，政府管制造成这样一种环境，它诱导单个雇员不使用可以获得的决策技能。他们是利用不足，没有开动全部"汽缸"。它们没有以直接闲暇换取直接努力，什么都未交换，相反这是某种东西未被充分利用的情况。因此，它不能看作是减小闲暇"价格"导致(假定是理性行为)闲暇数量增加。

从X效率理论观点来看，对价格与数量的这种相当传统的解释有两个难处。第一，全体雇员都对成本和X(低)效率有关系。但是，全体雇员都知道因政府管制造成的闲暇价格差异吗？或者，政府管制只影响部分雇员——譬如说，上层经理人员。第二，让我们假设情况为后者，政府管制允许上层经理人员的行为有更大自由处置权。特别是，他们的行为很少受应该达到X效率需要的约束。我们甚至假设，增加自由处置权就是增加他们的效用。但是很有可能，增加自由处置权程度，会使那个人的行为更加任性，从而降低他们下属人员的效用。无法先验假设，一批人增加的效用超过另一批人减少的效用。换言之，政府管制可能造成或增加X低效率。它这样做没有增加效用，或者因为它不增加效用。

我们用图5-7说明这个情况。垄断企业和竞争企业要价分别为P_m和P_c；产出率分别为Q_m和Q_c；成本分别为C_m和C_c。垄断企业利润——面积1——用P_mABP_m'度量，较高的工

资——面积 2——用$P_m B' B'' P_m'$度量。按新古典理论的观点，面积3和面积 4——$P_m'' B'' B P_c$——的存在，代表闲暇价格下降引起的管制的效用收益。根据定义，不可能存在X低效率。但是，却存在其他两个可能性。第一，当一些人得到的效用超过其他人的负效用时，就存在一定的净效用。在这种情况下，面积 3 代表效用，而面积 4 代表X低效率。第二，没有〔净〕效用产生，以致面积 3 和面积 4 都代表X低效率。

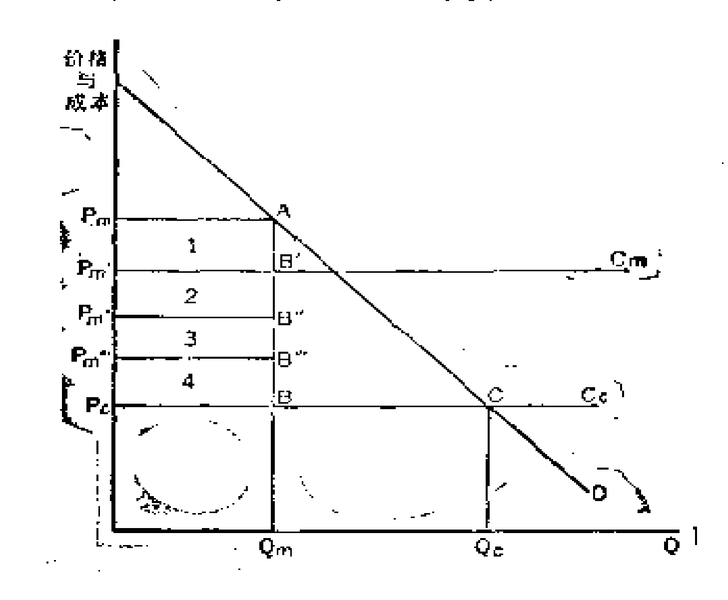


图 5-7　垄断力量造成的较高成本划分：利润、工资、效用和X低效率

A.5.3　报酬率管制的X低效率和配置低效率效应

让我们假设，政府通过报酬率管制来管制公用事业。因此，企业允许的收益要求相等于经营费用加上当期折旧，再加上企业允许的报酬率乘上资本价值（“费率基础”）的积。这个由包括费率基础在内的若干因素决定收益要求的程序导致所谓的“资本偏向”、“过度资本化”，简而言之就是配置低效率。

图 5-8 中，我们有一条等产量曲线 Q_1 和代表市场价格的等成本曲线 R_1。因此，追求成本极小化的公用事业企业将沿射线 Z_1，在 A 点以劳动 L_1 和资本 K_1 进行生产。但是，报酬率管制造成资本偏向，使资本价格实际上相对于劳动价格“下降”，从而与斜率较小的等成本曲线（例如 R_2）相关联。这就是说，管制会鼓励企业更加资本密集化。为什么不是呢？较大的费率基础或资本存量将导致较大的收益，主要取决于许可的报酬率。根据这些“新的”投入价格，企业就在等成本曲线 R_2 的约束下进行生产，使用劳动 L_2 和资本 K_2，在射线 Z_2 上的 B 点进行生产。企业在资本偏向既定的情况下使其成本达到极大化。

可是，从造成等成本曲线 R_1 的市场价格角度来看，企业并未使它的成本达到极小化。从市场价格看，企业过于资本密集了。显然，资本的市场价格高于“管制的”价格。假使企业面对管制价格，很“自然”它将更多地使用资本。因此，单

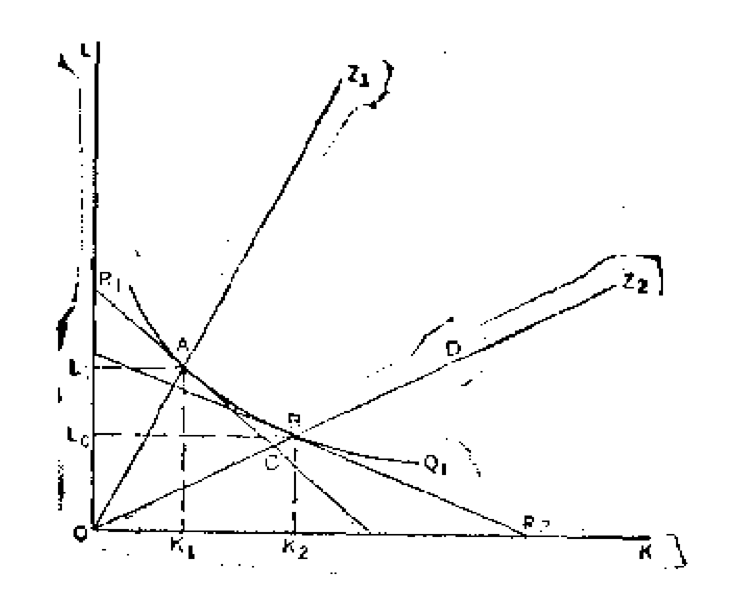


图 5－8　报酬率管制，配置价格和 X 低效率

从市场价格的角度看，企业的资本/劳动比率就显得太高了。因此，从市场价格的角度看，企业可以通过较多地使用较便宜的投入(劳动)来降低它的成本。

在图 5-8 中，报酬率管制引起的较高成本用 OB/OC 衡量。点 B 是重要的，因为这点代表了生产的等成本曲线。点 C 的使用是因为它与点 A 在同一条等成本曲线上，因而代表相等的成本。其次应注意的是，点 C 比点 B 更接近原点，因此点 C 代表比点 B 低的成本。换个角度说，因为点 C 比点 B 更靠近原点，所以点 C 与代表较低产出水平的等产量线相交。这就是说，给定市场价格和相对较高的资本价格，使用相对较高资本/劳动比率的企业，能够购买较少的投入，从而生产较少的产量。如图中 Q_0 所示。

但是，这些较高的成本并不是企业低效率，因为企业可以在点 A 和点 B 上使其成本达到极小化。这些较高的成本归因于投入价格的扭曲，在我们所举的例子中引起投入价格扭曲的是报酬率管制。因此，这些较高的成本是配置低效率的一种形式。在这里配置低效率由 OB/OC 衡量。此外，就企业取得允许的报酬率能够弥补它的较高成本，以及不管企业成本多少管制者都会维护它们而言，企业还是获得不必使它们的成本达到极小化的激励。在图 5-8 中，不论从市场价格还是从管制价格的角度来看，生产 Q_1 的非极小化成本如点 D 所示。OD/OB 就是 X 低效率的一种衡量尺度。X 低效率并非投入价格扭曲的结果，而是缺乏成本意识，对可获得的技能利用不足的结果。因此，由报酬率管制引起的总的低效率是 OD/OC。正如已经指出的那样，经济学理论只集中于研究配置低效率。

使配置低效率极小化的拉姆齐(Ramsey)定价法

有这样两个问题。第一，如果报酬率管制造成配置低效率，那末为什么还要用这种方式进行管制呢？在提出第二个问题之前，我先回答这第一个问题。人们常常假定，公用事业企业是“自然”垄断企业，也就是说，它们的成本函数呈连续的负斜率，表示生产处于连续的规模经济中，或者该行业是成本递减行业。借助图 5-8，我们把这种情况用图 5-9 来说明。

我们首先假设企业能够获得的最好绩效是在图 5-8 和图 5-9 所示的点 A 上。图 5-8 的点 A 表明企业按市场价格进行生产，产出率 Q_1 达到了成本极小化。这些极小化成本是企业“整体”极小化成本曲线上的一点。在图 5-9 中，企业在这条整体极小化成本曲线 MC_m 上进行生产，并满足 $P=MC_m$。因

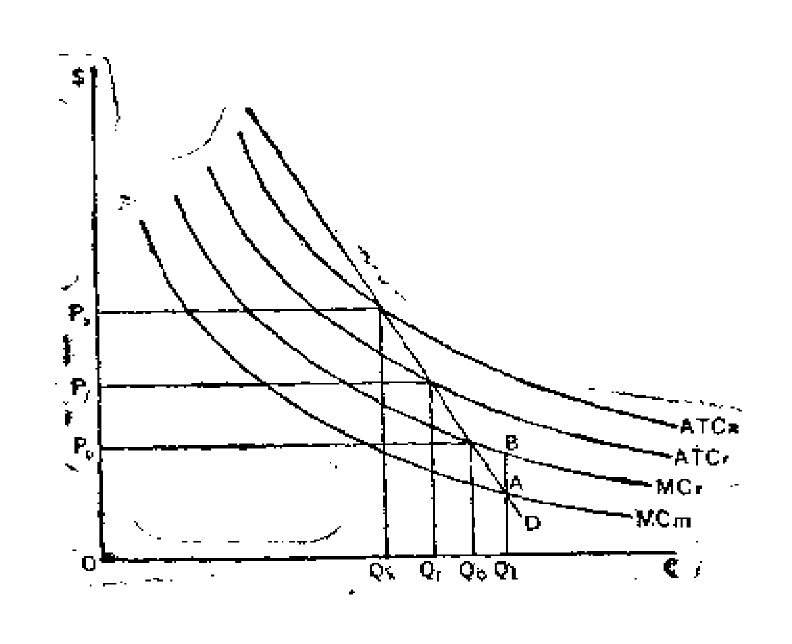


图 5－9　管制，拉姆齐定价，配置低效率和X低效率

此，可以假定点A是无法达到的理想点，因为不能预期不受管制的垄断企业会满足图 5-8 或图 5-9 所示的点A条件。

于是政府试图通过报酬率管制把企业推向点A。但是，恰如我们已经了解的，报酬率管制会提高生产 Q_1 的成本，使得受管制企业的成本曲线向上移动到 MC_r。在图 5-8 和图 5-9 中，受管制企业将在点 B 上生产 Q_1。

假定管制委员会并不清楚地知道企业技术上能达到的极小成本。委员会甚至企业本身，很可能会假定这些较高成本是它们真实成本函数的一部分。人们认为，这个成本函数 MC_r 是整体极小的，而事实上它却不是。但是，假定政府把价格定在 P_0 上，使得企业生产 Q_0。在 Q_0，$P=MC_r$。可是在 MC_r 上，企业获得的是负利润。

于是委员会管制企业，以使它能获取"公平的"报酬率，这里的"公平"允许企业弥补它的经营费用、折旧和它们费率基础上的公平报酬率。在这种情况下，企业将生产 Q_r，要价P_r，但是社会得到的仍少于社会最优产出率(在Q_r 上，$P_r>MC_r$)。因此，对第一个问题的回答是，我们管制被认为是自然垄断的企业，因为：(1)我们希望达到整体最优，而相信舍此便别无他法；(2) 假定企业是成本递减行业的一部分，如果企业在 $P=MC_r$ 处进行生产，它就将得到负利润。

现在，我们看到有两种形式的配置低效率是这个管制过程造成的：(1)社会得到的是低于社会最优的产出率；(2)企业是过度资本化的。因此，第二个问题是，我们怎样才能使配置低效率达到极小，也就是避免偏离社会最优产出率，同时允许企业获得公平的报酬率。幸运的是，经济学理论对此有

了回答：拉姆齐定价法(Ramsey Pricing)[1]。

考虑一个公用电力企业，它有三类顾客：居民、商业和工业。拉姆齐定价法这样规定，每一类顾客的电力产出以及对每一类顾客索要的价格必须使其价格偏离边际成本（MC）的百分数与该类顾客需求的价格弹性成反比例。因此，对具有相同弹性的顾客将按(供应该类顾客的)边际成本以相同百分比加价索要。对具有较小(较大)弹性的顾客索要的价格偏离边际成本也就较高(较低)。满足这个条件的价格就是拉姆齐价格。拉姆齐定价使对于社会最优产出率的偏离达到极小，所以它能使这种类型的配置低效率达到极小。在图5-9中，假定价格 P_r 是企业对它三类顾客索要的某个平均拉姆齐价格，它使得企业得自这三类顾客的平均利润为零，对 $P=MC_r$ 的偏离达到极小。

A.5.4 报酬率管制的替代办法，配置和X低效率

导言

在图5-9中，我们已说明，允许企业有公平的报酬率会造成配置低效率，但是，使用拉姆齐定价法能使这种配置低效率降至极小。报酬率管制也会诱发X低效率，这里所说的X低效率由企业的实际平均成本曲线（ATC_x）超出我们假定的企业技术上极小化整体成本曲线 MC^* 的部分表示。允许企业用X低效率来取得公平报酬率就要求价格为 P_x，这个价格反过来导致产出为 Q_x。

---

[1] 拉姆齐定价法最早在拉姆齐(1927年)的著述中有论述，以后由鲍莫尔和布拉德福德(Baumol and Bradford,1970年)"重新发现"。

这里出现这样几个问题。第一，有没有取代报酬率管制的办法？第二，X 低效率重要吗？第三，要是重要，怎样才能使它达到极小化？第二个问题主要在下一章里回答，在那里我将提出受管制行业尤其是在公用电力事业中存在X低效率的经验证据（关于经验证据的以后两章还将论述对于受管制行业的一些经验研究）。如果X低效率是重要的——我相信经验研究表明的就是这样——那么，怎样才能使它极小化呢？

已提出了若干替代报酬率管制和使受管制企业X低效率极小化的办法。克鲁和克莱因多弗（Crew and Kleindorfer，1986年）对几个这样的建议作了综合概述，下面我就根据他们的框架进行阐述。

激励管制

曾经提出过几种不同的激励管制办法。第一种是苏迪特(Sudit,1979年)建议的，由管制委员会确定一个允许的年度价格增加百分率，这个百分率由预期的生产要素价格上涨和相抵消的生产率提高所决定。企业因此受到激励使生产率增长的速度超过委员会“允许”的速度，因为假使其他条件不变，生产率的较快速度增长，会转化为成本的降低和（企业能保留的）利润的提高。根据图5-9，企业得到激励要将它的成本曲线从 ATC_x 降低到 MC^*，或者至少降到 MC_r。克鲁、克莱因多弗和苏迪特(1979年)假设，委员会不可能像企业那么精确地预料价格和生产率，它们试图向企业提供激励，以促使其向委员会透露这种信息。

洛布和马加特(Loeb and Magat,1979年)建议，只要企业把价格降到管制价格($P_r=ATC_r$)以下，并趋近社会最优产出率

Q_c,就给予企业补贴。激励应该使企业利润在 Q_c 实现极大化。虽然这将减少配置低效率，但是如果激励合乎理想，那它就能导致向产出率 Q_c 的移动，在这个产出率上，配置低效率和X低效率都将等于零。克罗斯(1970年)使用这样一种激励定价公式：$P=a+b(ATC_y)$，这里 ATC_y 是委员会估计的企业平均成本，a 是价格下限，b 是"分享率"。例如，如果委员会规定 $b=0.4, a=6$ 美元，$ATC_y=10$ 美元，那末企业的价格就为10美元。如果企业能把它的平均成本降低到 8 美元，那末它面临的价格就是9.20美元(6美元+0.4×8美元)。企业因此能够保留 1.20 美元的成本节约或X低效率收益。另一方面，如果企业成本增加到 11 美元，那它就将面对 10.40 美元(6美元+0.4×11美元)的价格。假定企业希望得到较高的利润，那么激励就应放在降低成本并使之接近 MC^* 上。

特许权招标

每隔一段时间，就会有一次公开的招标，以决定谁有权经营这个受保护的垄断企业。每个投标者必须确定未来一个时期里他(她)将索要的价格。如果在与这些确定价格相关联的剩余中利得的贴现值超过现在经营的价值，那末特许权就归于这位中标的投标者。使用特许权招标的建议是由登姆塞茨(Demsetz,1968年)提出的。就X低效率而言，预期的结果将强迫决策者把成本保持在必要的低水平内，即保持在可以继续控制特许权的水平内。虽然不一定就要降低到 MC^*，但只要有足够多的投标者，那么就可以预期成本至少会接近 MC^* 的水平。

放松管制

关于放松管制的一种观点假定，因为从放松管制中得不到什么好处，所以放松管制没有什么意思。麦考密克(McCormick)、舒加特(Shughart)和托利森(Tollison)都持这种看法。他们的立场似乎基于这样两个论断：第一，管制的社会成本"绝大部分"是原来创造管制的寻租支出；第二，这些建立垄断之前的寻租支出等于全部未来垄断(受管制的)利润的现值。换言之，全部租金都消耗在寻租活动中；受管制的垄断企业的预期利润等于零。这个主张与塔洛克(Tullock)的这样一个看法相似，即超过最低水平的全部成本——图 2－7 中的面积 P_mABP_c——都归因于像理性投资一样的寻租支出。当然，这些作者在得出这些结论的时候，都是假定不存在管制引起的 X 效率损失的。情况倘是这样，较高的成本可以假定完全来自寻租支出。

这些作者还运用供求曲线说明这一点。创造管制的寻租支出的影响会使受管制商品的边际成本曲线向上移动，造成一个以低产出和高价格为特征的新均衡位置。如图 5－10 所示，竞争性的行业按价格 P_c 生产 Q_c。受管制的垄断企业按价格 P_m 生产 Q_m。但是，回想一下他们的假设：管制的较高成本完全归因于寻租。这种寻租支出使边际成本曲线永久性地上移到 MC_r。这些支出是一种社会成本，它们表明资源离开生产用作寻租。这样做的话，经济将越来越贫困。由寻租造成的较高成本代表那些不能够用随后的放松管制来"降低"的成本。根据图 5－10，恢复竞争后经济不会回到 Q_c，而只能回到 Q_c'，其差额就是已被用于寻租的资源。放松管制造成的福利增加为图中三角形面积 ABD。很清楚，从放松管制得到

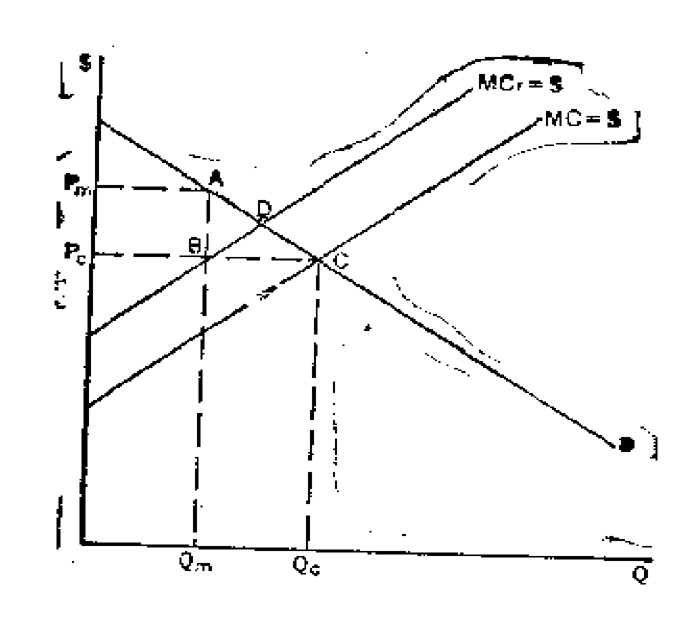


图 5－10　对放松管制没有兴趣，因为假定放松管制对X低效率没有影响

的这些好处要小于图中由三角形面积ABC表示的初始的净经济损失。

因为放松管制的得益小于成本，所以在这些作者看来，听不到公众普遍要求放松管制的呼声也就不足为奇了。这些作者宁愿对管制的利益与成本作“简单的”经济学争论，而不愿像有些人提出的那样，就民众对经济相对无知或消费者利益集团软弱无力这样的问题进行争论。在本书的分析中，这是一个重要的问题，因为在这种情况和第8章将提到的另外一些情况下，简单并不一定好，所以没有必要作这样的选择。在简单解释假定不存在各种形式的低效率就使这种简单解释一开始值得怀疑的情况下，这种说法尤其确切。

例如，如果放松管制有X效率方面的好处，那些相信他们能够以比受管制垄断企业更低的成本提供商品与劳务的生产厂商和那些相信放松管制能够降低价格的消费者就会对放

松管制感到兴趣。美国电话电报公司耗尽它所有的租金这一事实，并不能影响斯普林特公司（Sprint）进入电信行业的决策。事实上，很可能正是因为看到原有企业中存在的X低效率，才使其他企业寻求进入。换言之，假使其他条件不变，X低效率的存在会使放松管制的利益变大，因为X低效率成本是不断提高的。

一个相关的问题是所有的租金是否都已耗尽。虽然关于寻租的早期论文——塔洛克（1967年）和波斯纳（1975年）——都假定所有的租金都消失，但后来的论文却认为它们不会完全消失，因为存在这样一些因素，诸如不同的开办成本［罗杰森（Rogerson，1982年）］和风险避免［希尔曼和卡茨（Hillman and Katz，1984年）］，或者因为对企业的转移并不是寻租支出的线性函数（塔洛克，1980年）。此外，克鲁和罗利（Crew and Rowley，1986年）补充了有限理性这样一个原因。克鲁和罗利还指出，麦考密克及其同事们的理论暗示，管制持续存在是因为管制市场是均衡的。但是，正如克鲁和罗利所说，管制的持续是由市场和政治共同决定的。这里重要的是，管制市场均衡的假设和管制企业的较高成本完全由寻租支出造成的假设，导致麦考密克及其同事们得出这样一个结论：放松管制的潜在收益很小。我相信，麦考密克及其同事们的论文是这些假设为什么重要的一个例子。

另一种看法是，放松管制是取得效率收益的重要潜在来源。放松管制能采取什么形式和预期能取得什么收益呢？放松管制能采取三种形式：放松对垄断的管制，自我管制和虚假竞争。在放松垄断管制方面，管制委员会和企业的独占特许权将被取消。因此，进入是合法的。由于竞争，可以期望

X 效率和配置效率都将增加。自我管制就是医师、牙医和法律职业实行的那种办法。国家负责批准执照和进行监督，如果这些职业对消费者不负责任，国家随时都会进行干预。由于报酬率管制限制了企业的利润，所以，可以预期自我管制会增加X效率。虚假竞争指的是管制与竞争并存。美国电信行业在美国电话电报公司的专营权被剥夺后的情况就是虚假竞争的一个例子。随着竞争增加，可以预期X效率也会增加①。

① 根据芝加哥贝尔公司雇员的样本，把心理变量和经济变量结合在一起进行放松管制的实例研究，可参见S·马迪(Maddi)和S·科巴萨(Kobasa)1984年的论文。



经验证据：受管制的行业

6.1 导言

此刻，我们已经介绍了新古典主义理论（一种公认的有一些正当理由的机械论的论述）、X效率理论和几个别的偏离新古典主义正统理论的主张。这些偏离正统观点的主张（包括X效率理论）在某些方面具有共同的要素，而在某些方面每种主张又都是独特的。我们还已提出X效率理论所牵涉的各种含义。因此，现在是阐述与X效率理论这些含义相一致的经验证据的时候了。关于这种证据有几点应该记住。第一，证据可分为三种类型：（1）关于对受管制行业的研究；（2）关于对企业所有制形式的研究；（3）关于对市场结构的研究。第二，这些研究中的任何一个都可以探究这些类型的两个或更多方面。每个研究之所以列入某个类型，是因为看来这个类型是该研究的重心所在。第三，在此后三章中提到的大多数研究都明确地讨论了X效率理论。读者还需要注意到那些不是明确地讨论X效率理论的研究。之所以要包括后一种研究，是因为只要其结果与X效率理论的结果相一致，这样做

看来就是有道理的。尽管这样，我尽量做到把这些类型的研究数量减到最小。因此除非另有说明，每篇论文都明确地讨论了X效率理论如何与其本身的研究相联系的问题。本章里我们要考察受管制行业中X效率的研究，包括电力与煤气公用事业、银行业、航空公司、为学校提供交通运送服务的单位和地方图书馆。我们先从电力和煤气公用事业开始。

6.2 电力公用事业

6.2.1 普里莫克斯

电力公用事业中的成本和竞争

如果经济或技术条件允许一个行业只能有一家有效率的企业，那末那家企业就被称为“自然垄断企业”。导致存在自然垄断企业的一个条件就是在整个行业需求范围内具有正的规模经济效应。这个说法导致了20世纪初以来出现的争论：即政府应该控制垄断特许权，并把它给予电力、煤气和电话服务的生产者——公用事业。

以规模经济为基础的自然垄断的理由，依赖于这样一个假设：X(低)效率是零。也就是说，较大的(垄断)企业通过规模经济产生利益，并由于它的垄断地位，不会产生任何过多的成本。可是，这些企业面前没有竞争的压力，结果是它们的成本可能超过技术上的最小成本。如果说生产成本不受市场结构的影响，那末公用事业的垄断化便成为比较合适的公共政策。另一方面，如果垄断将成本提高到超过竞争的行业水平，那末垄断化的净效果，就要取决于规模经济（它降

低垄断成本）和X效率（它提高垄断企业的成本）之间的相对力量。关于规模经济和X效率的相对效果，沃尔特·普里莫克斯（Walter Primeaux，1977年）在他对美国电力公用企业的研究曾加以衡量。

普里莫克斯使用联邦动力委员会公布的资料比较了1964—1968年期间公众控股的电力公用事业的独家垄断企业样本中的平均总成本（以扣除税收和相当于税收的金额后的总成本除以用千瓦小时计算的年供电量）和公众控股的电力公用事业的两家垄断企业样本中的平均总成本。（在调查期间，美国大约有49个地区的电力使用者可以从两家企业选择使用任一家电力。在大多数情况下，两家垄断企业与一家私人控股的企业竞争。）尤其是，比较是在面对竞争的公众控股企业的成本与在一个特定州的"垄断城市"里的公众控股企业的成本之间进行的，并且尽可能按照大小规模和动力来源（蒸汽、水力或燃气机）分别进行比较。例如，把一个特定州里的"小型"蒸汽发电的独家垄断企业与"小型"蒸汽发电的两家垄断企业进行比较。

使用普通最小二乘法（OLS），普里莫克斯控制了销售额、发电能力利用水平、燃料成本、每个客户的耗电量、市场密度、内燃发电、12个州的虚拟变量、产品市场竞争的一个虚拟变量、以及销售与竞争之间相互作用等变量。普里莫克斯的资料的确表明独家垄断企业中随着产出增加单位成本会有所下降，也就是说，独家垄断的公众控股的公用事业中规模经济是有活力和起作用的。图6-1中，这个情况用向下倾斜的独家垄断企业长期平均总成本曲线（MM′）表示。但是，对于两家垄断企业来说，随着产出增加，长期平均成本DD′会

有微小的增加。

第二，根据样本中所有企业的平均产出率(样本平均数)，两家垄断企业的单位成本较之相比较的独家垄断企业低10.75%。在图6-1中，这个情况表示为这样一个事实，即在平均产出率为Q*时，两家垄断企业的成本(AQ*)低于独家垄断企业的成本(BQ*)10.75%（我要求读者想象AQ*比BQ*小10.75%）。第三，这种差距在较小企业之间还要大，也就是，人们相信小企业比大企业更多地受到竞争的"惩罚"，换言之，面对竞争的较小企业从增加的X效率那里得到的收益会比较大。第四，当每年产出率达到2.22亿千瓦小时(kwh)时，两家垄断企业与独家垄断企业的单位成本相等。第三点和第四点的结果表现为这样的事实，即在竞争——两家垄断——的环境中供应电力的成本优势随着产出增加到接近2.22亿千瓦小时后便减弱了，在这个产出率上，两种市场结构之间的长期平均成本是相同的。第五，到产出率超过2.22亿千

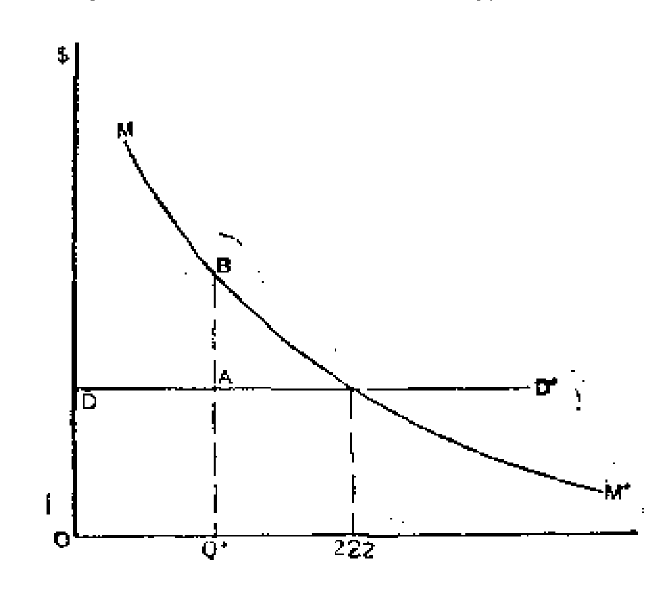


图6-1 规模经济和X效率：竞争在电力公用事业行业中的影响

瓦小时时，独家垄断企业的单位成本就较低。这第五点的结果表现为当产出率大于2.22亿千瓦小时时，DD′处在MM′之上。

对这些资料有一个解释，那就是规模经济的作用是存在的，但是在产出率低于2.22亿千瓦小时水平时，X效率的作用超过了规模经济的作用。反过来说，当产出率大于2.22亿千瓦小时时，规模经济的作用则超过X效率的作用。此外，这些资料还指出，较小企业的X效率效应较大。值得注意的是，在本世纪60年代初期，在3190家电力公用事业中（取样研究的有8%的公众控股企业和27%的私人控股企业），只有437家的年产出率超1亿千瓦小时。在这些年产出率超过1亿千瓦小时的437家企业中，也并非全部的产出都达到2.22亿千瓦小时。换言之，在那个时期研究的样本中，除了那些最大的企业，通过竞争获得的X效率的利益都超过通过垄断化获得的规模经济的利益。普里莫克斯的结论是，"出于政府制订政策的目的，单是X效率的存在就反映出对受管制行业（以自然垄断形式被进行管制）进行重新审查的重要要求。"

电力公用事业中生产能力利用率与竞争

对上面论述的普里莫克斯研究的一个批评是，可以断言垄断企业比（较小的）竞争企业有较大的生产能力。这可以推出，垄断企业的较高单位成本可能不是因为X低效率，而是由于维修较大量的过剩生产能力的需要。电力公用事业独家垄断企业的生产能力较之两家竞争的电力公用事业双头垄断企业的联合生产能力更大吗？普里莫克斯检验了这个认为独

家垄断企业有较大生产能力的假设。如果证明这个假设不对，那末正统的观点，即以过剩的生产能力来解释垄断企业成本劣势的说法就可以抛弃了。这点本身不足以证明X(低)效率的存在和重要性。但是，因为这个假设是用来作为"解释"在受管制的垄断企业中"不存在"X(低)效率的手段，所以驳倒这个生产能力假设将使X效率解释成为一种可行的方法。

普里莫克斯(1978年)使用联邦动力委员会公布的1964—1968年期间的资料检验了这个生产能力假设。他以研究成本时所采用的同样的比较方法研究了"独家垄断城市"和15个"两家垄断城市"的市政当局控股的电力公用事业的资料。这里他再次使用了普通最小二乘法，即对销售额、最大负荷条件、公用事业购入的能源量占其销售量的百分比、对不同顾客阶层的销售分布、各州的虚拟变量、以及假设企业是竞争性(两家垄断)时的虚拟变量的控制。

他的研究结果包括发现竞争对生产能力利用率没有影响。换言之，两家竞争的市政当局控股的电力公用事业的联合生产能力条件与独家垄断企业的生产能力条件并无不同。用生产能力来"解释"垄断企业中的较高成本和由此反对用X效率来解释较高成本，因而就显得与这次对生产能力利用率和市场结构研究的发现大相径庭。

"自然"垄断是自然的吗？他的最新著作《电力公用事业的直接竞争：自然垄断的神话》

普里莫克斯(1986年)报告了与主张为何把公用事业视作一种自然垄断企业的许多论断不一致的证据。因为自然垄断论点是管制的公共利益理论的一部分，所以驳斥前者也含有

(部分)驳斥后者的意思。兹将他的若干结论报道于下。

第一，从19世纪中叶起，就假定规模经济对一家电力公用事业极端重要，以致使竞争成为低效率的。但是如业已讨论的那样，普里莫克斯的研究表明，除了最大几家电力公用事业企业外，竞争(增加X效率)带来的利益要超过规模经济的利益。

第二，自然垄断，包括电力公用事业的自然垄断，据说有相对高的固定成本，这允许它们的长期平均成本落在产出的整个范围里。普里莫克斯于1968年比较美国几个行业的资本-产出(K/O)比率来检验这个说法，并使用了《商业概况》月刊中的数据。所使用的资本计量法考虑到了折旧，因为资本的使用寿命在各行业间不同，在行业内部也不一样。因此，16个行业或部门的每一个资本-产出比率是以总产量或销售量除以折旧来计算的。行业间的比率从0.232(农业和矿业)到0.53(建筑承包业)不等，16个行业的平均比率为0.085。

在自然垄断企业中间，其中比率为0.142的全为交通通讯行业，其中比率为0.145的有电话、电报和相关服务业以及运输业，铁路的比率为0.151。这些比率不但被上面列举的行业超过，而且被农业(0.224)超过。此外，所有金融业、保险业和房地产业的比率(0.151)与铁路的比率相等。最后，密苏里公用事业公司的资本-产出比率为0.079，这个比率使该公司低于除了三个行业以外的所有行业，并比16个行业的平均数字还低。普里莫克斯的结论是，公用事业的固定成本不是它据有自然垄断地位的正当理由。

第三，据说自然垄断条件会使竞争变得不可能。但是普里莫克斯指出，两家垄断企业之间的竞争在密苏里地区已存

在四五十年。第四，据信电力公用事业间的竞争将把价格提高到超过垄断条件下的水平。这个说法的根据是认为竞争性公用事业的成本高于垄断条件下企业成本的观点。普里莫克斯则指出，价格在竞争条件下较低，最小的差距为16%(500至750千瓦小时那组)，平均差距为33%。

另一个可能令人吃惊的结果是普里莫克斯报道的关于消费者了解价格差距的程度。在一次用电话对赛克斯顿、密苏里31个家庭的任意抽样调查中，普里莫克斯报道说，有22个家庭可以从市里得到两家公用事业公司的服务。这22个家庭中有17家不知道哪家企业收费较低，只有5家能准确地回答这个问题。22个家庭中还有5家甚至不知有多少企业出售电力。

总的说来，普里莫克斯的研究结果表明，电力公用事业的特点不完全符合自然垄断的理论。这并不是说电力公用事业不应受管制。但是，假使它们应受管制，作出管制决定的理由不应是它们是作为自然垄断者而存在的。最后，普里莫克斯的结论令人怀疑管制的公共利益理论，而自然垄断观点就是该理论的一部分。对它的怀疑为X效率假设提供了支持。

6.2.2 史蒂文森

潜在的权衡不但存在于X效率与规模经济之间，也存在于X效率与“范围经济”(Economies of Scope，又称“联合经济”Economies of Combination) 之间。当一个企业生产二种产品其成本比两家单独的企业生产它们便宜时，范围经济便出

现了。范围经济的一个事例是一家公用事业同时出售电和天然气供家庭取热之用。这些"联合"公用事业就这样在"家庭取热"市场出售两种主要的竞争产品。(据信这个市场中的范围经济主要存在于企业的分配和一般管理部门，也就是在开票、抄表和两种产品分配管理的方面。)这种做法使企业面对的竞争压力比它只生产两种产品中的一种要小。

但是，虽然范围经济有助于降低联合公用事业的成本，使之低于只生产电或天然气的公用事业——"单一"公用事业——的成本，但缺乏竞争压力可能通过X效率效应会使成本提高到超出单一公用事业的成本水平。联合公用事业与单一公用事业的相对成本将依靠范围经济和X效率二者效应的相对力量来决定。

罗德尼·史蒂文森(Rodney Stevenson,1982年)的研究，考察了79家美国公用事业样本中生产电力的成本(其中25家为联合公用事业，54家为单一公用事业)。他把联合公用事业限定为其总收入中至少有15%来自天然气的企业，而单一公用事业是其总收入中来自天然气的比例不足5%的企业。在比较发电成本中，史蒂文森不能发现范围经济和X效率的相对影响。但是他能够测出市场结构——两家垄断与独家垄断——对X效率(成本)的效应。他的资料来自联邦动力委员会和全国公用事业管制专员协会。

在估计1970年和1972年所有79家公用事业的对数成本函数时，史蒂文森把投入价格、产出、工厂生产能力利用率和东北部、西北部、西部和南部的地区虚拟变量包含在内作为自变量。使用1970年的数据，史蒂文森报道说，在样本平均值(自变量的平均值)上，单一公用事业的平均发电成本

比联合公用事业低6.1%。使用1972年的数据，他报道说单一公用事业的平均发电成本比联合公用事业低8.5%。由此可见，当企业面对较大竞争压力时，它的“静态”效率较高。

史蒂文森接着测试竞争压力的“动态”效率效应。动态效率是指企业一段时间里减少成本的能力(它的实际表现)。一段时间能较大地降低成本就证明有较大的动态效率。为了做到这点，他比较了1964年和1972年单一公用事业和联合公用事业的成本函数，以期从“时间趋势”中分离出（上面列举的）自变量的效应。这个方法表明，单一公用事业在这段时间里减小成本的速度的确比联合公用事业快。(不是全部试验都显示动态效率存在有统计显著性的差异。)换言之，竞争压力不但导致在时间的任何点上降低成本，而且导致在整段时间里成本函数有较大的向下移动。

史蒂文森得出结论说，竞争确实影响生产成本(X效率)。“虽然现存的公用事业和管制委员会（联邦通讯委员会例外）历来都证明是厌恶竞争安排的，可是新企业的闯入和企业间的竞争可能对于推进公用事业经营活动的改善是一种非常有用的管制工具。”

6.2.3 戈洛普和卡尔松

在70年代早期，燃料价格开始飞速上涨。电力公用事业担心它们的收入抵不过螺旋上升的成本，于是向它们的管制委员会申请提高电力价格。典型的答复是允许公用事业只能针对较高的燃料价格自动提高电价，也就是它们不可以为较高的劳动成本、资本成本或非燃料原材料成本自动提高价格。

这种自动调整机制称为 FAM——燃料调整机制。在图 6-2 中可以看出 FAM 的潜在效应。在纵坐标上，我们计量非燃料投入，而燃料在横坐标上予以计量。反映市场价格——非 FAM 范围——的等成本曲线为 NF。因而成本极小化的公用事业将使用OF_1的燃料与 ON_1的非燃料在点 A 上生产产出 Q。假使燃料调整机制允许公用事业只能弥补较高的燃料成本，成本极小化的公用事业就得按照好像燃料价格已经下降那样经营。换言之，燃料调整机制鼓励公用事业使用更加燃料密集的投入构成。等成本曲线因而向反时针方向移动至NF′，成本极小化企业将在点 B 上采用更加燃料密集的投入构成，使用 OF_3的燃料与 ON_3 的非燃料。假使燃料调整机制既定，点 B 代表一种成本极小的投入构成。但是，企业移离扩展线 OP，而趋向更加燃料密集的扩展线 OP_1。

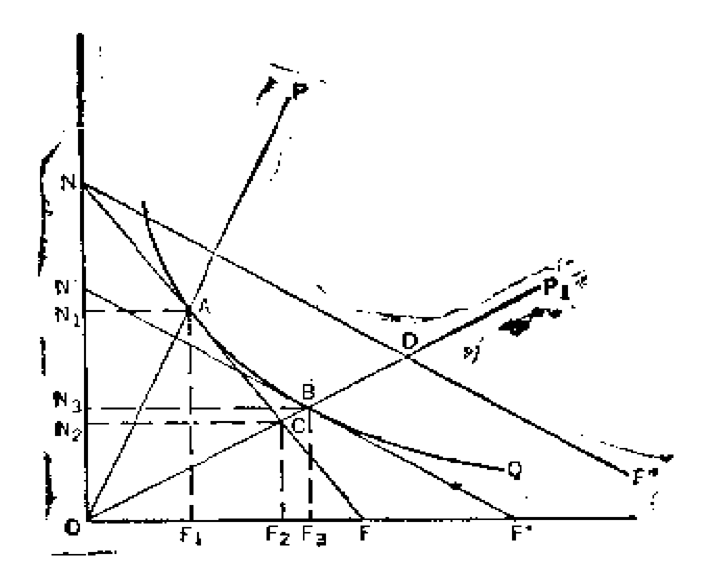


图 6-2　由燃料调整机制引起的配置价格和 X 低效率

以市场价格估价，这个新的投入构成显得过于燃料密集了。这点在图 6-2 中可以从这样的事实看出来，即面临市场

价格的企业——等成本线为 NF——如果它沿扩展线 OP_1 生产，不会生产产出率Q。那就是说，它的支出不允许它买进足够的“高”价燃料来生产产出率Q。假使既定市场价格和由等成本曲线NF给出的相关支出率，在扩展线 OP_1进行生产的企业只能达到明显代表一个较低产出率的点C。可见，为了生产产出率Q，企业必须增加它的支出或成本。

因为点A和点C在同一条等成本曲线上，它们代表相同的支出水平。要是企业使用市场价格，它将在成本 $OF_1(P_f)+ON_1(P_n)=OF_2(P_f)+ON_2(P_n)$ 的情况下生产产出率Q，这里 P_f 是燃料价格，P_n 是非燃料价格（所有非燃料投入的一个加权平均数）。但是在点B上，企业支出 $OF_3(P_f)+ON_3(P_n)$。这个支出率显然大于点C所代表的支出率。在这种情况下，它们受管制的鼓励去这样做。即使情况可能如此，以市场价格来估计，它们为每一个产出单位使用了过多的燃料和不足的非燃料。也就是说，它们属于价格配置低效率。价格配置低效率可根据距离OB/OC来衡量。当然，假如存在燃料调整机制，点B是成本极小点。

除了这个配置效应外，燃料调整机制还可能导致X低效率。因为燃料调整机制允许企业自动弥补较高的燃料成本，企业减小了成本极小化的压力，即X效率。企业不但会改变它们的投入构成多用燃料，而且会使用超过必要的燃料量。燃料调整机制对X效率的效应表现为企业在点D上生产产出Q。根据“燃料调整机制价格”估价，企业使用了正确的投入构成，但是会为生产产出率Q使用过多的投入。点D既不在市场价格NF上，也不在燃料调整机制价格NF′上，它是成本极小化的一个点。因而可以 OD/OB 的比率衡量X低效率。

而且燃料调整机制引起的总低效率可用OD/OC来衡量。

戈洛普和卡尔松(Gallop and Karlsson，1978年)测量了美国 105 家私人控股电力公用事业企业的燃料调整机制的影响。他们把燃料调整机制定义为由燃料调整机制所弥补成本的百分比。一般来说，公用事业委员会为燃料价格规定了一个"基数范围"(公用事业的燃料成本)，这个范围是根据对市场燃料价格的估计确定的。当市场价格超出这个基数时，公用事业能够弥补这种较高成本。当市场价格落在这个范围之内时，燃料调整机制则不起作用。假使市场价格在范围以下，公用事业必须进行"赊欠偿付"。测量燃料调整机制的效应需要估计燃料调整机制对企业生产成本(发电成本)的净效应。要做到这点，就要估计公用事业的成本函数，它以投入价格、产出和燃料调整机制作为自变量。对美国东北部地区、产煤带地区和墨西哥湾沿岸各州地区公用事业 1970 年和 1972 年的资料进行分析就使用这个方法。正是在这三年期间，燃料调整机制的使用有巨大的增加。

据这两位作者报告，只有 1972 年在产煤带地区的公用事业中燃料调整机制才导致配置低效率。关于燃料调整机制是否导致X低效率，两位作者报告了二个具有显著效应的事例，一个是1970年东北部地区和产煤带地区的公用事业，另一个是1972年东北部地区的公用事业。全面来看，他们的报告指明，燃料调整机制增加 1%，即可补偿的燃料成本增加1%，会导致总发电成本增加不到1%的十分之一。这个成本弹性数字似乎无论如何微小，燃料调整机制的效应也是显著的。例如，东北部地区或产煤带地区1970年的燃料调整机制没有产生X低效率，当时这两个地区的燃料调整机制值分别

为 0.39 和 0.51。此后东北部地区的燃料调整机制值为 0.62（1971年）和 0.82（1972年）；产煤带地区的燃料调整机制值为 0.57（1971年）。如果 1970 年的燃料调整机制值一直保持下去，东北地区的总发电成本 1971 年将下降 2%，1972 年将下降12%；1971年产煤带地区的总发电成本将下降0.4%。但是，如果燃料调整机制值为1.0%（完全成本补偿），东北部地区1971年的成本将增加 10%，1972年为 3%，产煤带地区1971年的成本将增加5%。

作者在墨西哥湾地区没有找到任何燃料调整机制引发的X 低效率，那个地区超过 70%的销售量服从燃料调整机制，而燃料调整机制值平均为73%。两位作者解释这个现象时指出，墨西哥湾沿岸各州的公用事业委员会较之东北地区或产煤带地区的公用事业委员会更频繁地调整燃料调整机制的数值。因而，墨西哥湾沿岸各州的公用事业委员会能够更好地保持基数价格接近市场价格。墨西哥湾沿岸各州的公用事业委员会比它们在东北部地区和煤带地区的同行面临较大的风险在于，要末是燃料调整机制不起作用，要末是受人指责它造成赊欠偿付：这两种类型的风险都鼓励X 效率行为。因此这两位作者得出结论说，"继续监督燃料条款规定可能是阻止低效率行为的最有效手段，而与此同时允许公用事业在迅速通货膨胀时期很快弥补不断增加的投入成本。"因此，人们可能认为墨西哥湾地区各州的公用事业委员会在管理企业时，对它们的行为（成本）较为负责。在戈洛普和卡尔松分析的样本中管制者的行为差异并不是罕见的。普里莫克斯（1985年）查阅某些文献时发现了同样现象，表明管制者为了不同（经济的和个人的）原因，常常倾向让公众控股的公用事业维持垄

断地位。

6.2.4 奥尔钦、凯塞尔和乔斯科：一些离题话

戈洛普和卡尔松的论文直接与乔斯科(Joskow, 1974年)的论文有关，间接与奥尔钦和凯塞尔（Alchian and Kessel, 1962年）的论文有关。我将努力说明，戈洛普和卡尔松以及乔斯科的论文能够用来"答复"当X效率理论运用于奥尔钦和凯塞尔隐含提出的那种类型的受管制企业时，有人提出的批评。

奥尔钦和凯塞尔

奥尔钦和凯塞尔的论文可以认为是一组关于我称之为"有效市场"(Efficient Markets)的一部分。有效市场假定：所有雇员——经理和其他人——都是效用极大化者；由于生产和家庭行为两个"部门"均由效用极大化者组成，在生产理论和家庭行为理论之间加以区分是人为的；而资本市场是完善的。

对于垄断的福利成本，奥尔钦和凯塞尔的处理与塔洛克相同，他们都假定，这些成本超过传统上对市场配置低效率测定的量。除这些配置损失外，奥尔钦和凯塞尔还认为，垄断企业造成的刚性价格是由墨守成规的经理人员操纵的，这些人连同别的雇员享受着轻松舒适的工作条件。作者们指出，这些结果是从真实世界观察而得，不是传统经济学理论所包含的，这个理论认为，不论竞争的还是垄断的企业，它们的经理人员都同样警觉和雄心勃勃，并以同等的精力追求利润。

但是，如果目标函数从利润或财富极大化改为效用极大化，这些其他结果就与传统理论相一致。简言之，这些其他结果不能合适地使用利润或财富极大化的模型来分析。

在利润或财富极大化模型中，竞争企业或者垄断企业的经理以同样干劲追求利润，他们同样警觉和雄心勃勃。可是，如果垄断企业的经理，在他们做不到利润极大化时面对较小被开除的风险，他们为什么要与竞争企业经理同样卖力工作呢？奥尔钦和凯塞尔的回答是，完善的资本市场迫使垄断企业经理与竞争企业里的同行以同样的态度工作。他们说，“……尽管在产品市场里没有竞争，那些能最有利地利用垄断权力的经理们将控制产品市场。资本市场的竞争会把垄断权分配给那些能够最有利地使用这种权力的人。因而，……只要存在自由的资本市场，产品市场没有竞争并不意味着垄断企业的经理们与竞争企业相比有素质上的差异。”相应地，一般而言，垄断企业不应该比竞争企业低效率。奥尔钦和凯塞尔将低效率定义为这样一种情况，即企业在没有非金钱（金钱）收入方面损失的情况下可赚取更多的金钱（非金钱）收入。

虽然有人说，经济学家没有能力决定一个人效用函数中自变量的值，但他们能够预言，任何商品的“价格”越低，它就被利用得越多。奥尔钦和凯塞尔比这更进一步，他们明确地说到极大化行为的意义。例如，一个人为了与朋友共事——不与他歧视的人在一起工作——而牺牲的收入是一种“平衡差别”（equalizing difference）。也就是说，在边际上此人牺牲的一笔收入（边际成本或供给曲线）其价值刚好等于不与那些人共事的利益（边际利益或需求曲线）。简言之，区别对待的理性消费者会选择一定数量的相互影响，直到“区别对待的

市场”达到均衡为止。

在效用极大化模型中，垄断企业的经理们追求金钱财富的劲头可能比竞争性企业的经理要小，但是二者在追求效用上的劲头是一样的。垄断企业经理们在追求金钱财富中之所以劲头不足，是因为他们会面对金钱财富与非金钱财富间的一系列不同的相对价格或交换比率，出现这种情形是因为垄断企业的“主要罪恶”是赚太多的钱。对受管制垄断企业来说，过多的利润会招来降低企业价格的压力。相反，只有企业能够证明它的利润低于“公平的利润率”，公用事业委员会才准许核定较高的价格。至于对不受管制的垄断企业来说，过多利润有招来反托拉斯法打击的忧虑。在两种情况下，非金钱财富的相对价格——由高于最小必要成本的额外所得——减小了，而理性行为意指需要较大数量的非金钱财富。非金钱财富以会产生效用的商品形式出现，如费用帐户、豪华办公室等等。因此，非利润极大化并不表示X低效率，而是表示“通过非金钱收益的有效效用极大化”。

乔斯科

奥尔钦和凯塞尔的方法以主张效用极大化和受管制企业几乎没有刺激达到X效率的理论（成本意识）来否定存在X低效率。乔斯科的论文指出为什么受管制企业具有比奥尔钦和凯塞尔指出的更多的刺激去达到X效率。

乔斯科对管制委员会的基本观察结果之一是，它们有兴趣尽量减少冲突，并采取一种“自己活也让别人活”的态度对待企业。就是说，管制委员会往往倾向于不改变它们的行为，除非它们工作的经济或政治环境对它们施加了压力。对管制

委员会的这种看法不同于通常的看法，后者认为管制委员会永远处于准备随时扑向实际利润率超过“公平”利润率的企业的过程中。

在乔斯科看来，管制委员会的命令——“公平”利润率和良好的质量——由委员会来解释，它们关心的是公平的下限而不是上限。也就是说，委员会更关心保证利润率能保障企业的适当资本，而不关心企业通过良好管理能赚得的最大利润限度。因而，乔斯科相信，管制委员会本身不很关心控制利润率，它们关心的是保持名义价格不致上涨。于是，只要企业不增加它的价格或减低它的价格，它可以增加它获得的利润率，实际上允许它赚取任何可能的利润率。

从这点看来，企业之所以引起管制机关的检查，是因为它希望提高它的价格和(或)调整价格结构。检查因而规定了最高价格，企业经营必须遵守限价，直到下一次检查意见听取会的召开。另一方面，降低价格通常无须正式召开意见听取会便会得到批准。因此，乔斯科的观点是，管制过程基本上是被动的。管制检查总是由企业引起的，管制委员会并不持续地监督规定的——列入目录的——企业不可超越的利润率。乔斯科还明确指出，在一直下降或长期保持正常平均成本的期间，受管制企业有一种进行“高效率生产”的刺激。以我们的术语说，这些企业有一种使它们具有X效率的刺激，因为只要正常价格保持在由管制机构规定的限额之下，企业就能保有它的利润。

作为他所说的管制检查过程被动性质的证明，乔斯科指出，为利润率召开检查意见听取会的次数在较高的通货膨胀时期较多(有时非常多)，这种时期有1949至1952年、1957

至1960年、1969至1974年。60年代后期通货膨胀的不断加剧，和环境保护主义者与其他消费者团体中间警觉性的增加向管制机关施加压力，足以使它把被动做法改变为较主动的做法。一个结果是管制委员会开始采取“凭经验决定的方法”或“标准工作程序”。凭经验决定的方法之一就是燃料调整机制。结论是，看来受管制企业的确有充分的刺激做到具有X效率(是不是全部雇员都有相同的刺激还不清楚)。但是，乔斯科的论文也指出，管制机关对企业采取自己活也让别人活的态度促进了X低效率的发生。最后，正如戈洛普和卡尔松所作的结论，管制机关的行为既不是相同的，也不总是站在有利于X效率一边。

6.2.5 霍拉斯和赫伦

我们回来谈谈比较直接的经验证据，霍拉斯和赫伦(Hollas and Herren，1982年)估计了电力公用事业市场结构(独家垄断和两家垄断)的配置和X(低)效率的效应。为进行这个估计，两位作者估计了197家公众控股企业的非生产边际成本，其中17家在1972年面临直接竞争。他们把非生产成本的范围限定为顾客帐户支出、销售额和一般管理支出。用含有一个虚拟变量的非生产总成本函数估计竞争性市场结构，他们报道说两家垄断的市场结构会为本地居住的客户供电降低其边际成本每千瓦小时0.184美分。第二，他们报道说，在这些电力公用事业样本中，竞争提高了固定的非生产成本。

第三，样本企业的平均产出为12360万千瓦小时。凡产

出率超出 11080 万千瓦小时的竞争性公用事业，其平均非生产成本比垄断性企业为低。因此，对于大部分样本企业来说，有竞争企业的平均非生产成本较低。为了比较配置低效率和X低效率的相对规模，这两位作者使用由30家垄断企业组成的子样本。使用了两种边际成本曲线，一种是它们的实际边际成本，另一种是如果企业面临产品市场竞争成本将如何变化的估计成本。估计的配置低效率和X低效率的不同组成如图6-3所示。

第一，区域F、I与J的和，代表要价(P_d)超过它的边际成本 $LRMC_d$（即P＞MC）的两家垄断企业的配置损失。如果两家垄断企业要价等于它的边际成本($LRMC_d$)，那末它将生产Q_c。但是独家垄断企业的边际成本($LRMC_m$)比两家垄断企业高。因而，要是独家垄断企业定价相等于它的边际成本，它将生产 Q_c'。在以上两种情况下，P＝LRMC。区域J因而代表由于减小产出和较高价格而形成的配置损失。（给定P＝MC，但$LRMC_d$＜$LRMC_m$。）

其次，假使两家垄断企业的价格定得超过$LRMC_d$，那末它生产Q_d，这个数量小于要是P_d＝$LRMC_m$时的产出，而根据后者它将生产 Q_c'。区域F因此代表配置损失，因为产出减小了，而P_d＞$LRMC_m$。反过来，区域I代表由较高成本引起的那一部分配置损失。

区域C是净损失——丧失的消费者剩余。这个损失是由于行业独家垄断造成P_m＞P_d和Q_m＜Q_d形成的。区域E是丧失的利润，这是产出降到低于两家垄断企业的产出率时形成的。区域H代表较高成本，但也是由于行业的独家垄断造成Q_m＜Q_d而产生的配置损失。最后，但并非不重要，区域G代

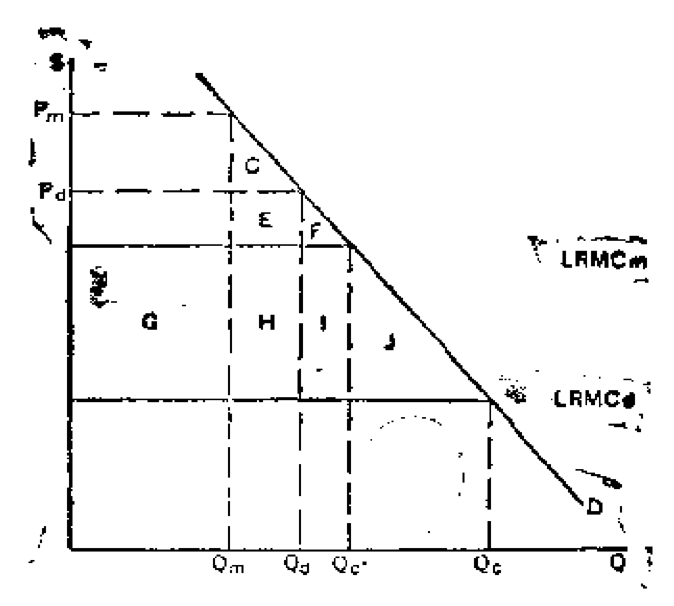


图 6·3 电力公用事业行业中由于缺乏竞争引起的配置低效率和X低效率的组成

表X低效率。

这两位作者作了几种估计，从假设P_d只略小于P_m，到假设$P_d = LRMC_d$。这里我将报道在$P_d = LRMC_m$情况下的估计数。这是正统微观经济学假定的情况：假使成本极小化$LRMC_d = LRMC_m$，假使竞争企业中存在配置效率$P_d = LRMC_d = LRMC_m$。在这个抽样调查中，当$P_d = LRMC_m$时，区域C——丧失的消费者剩余——等于样本中所有企业总收入的5%；区域(C+E+H)等于收入的7%，区域(F+I+J)等于收入的24%；区域G——X低效率——等于收入的8.9%左右。因此估计的总福利损失约占总收入的16.1%。如同他们的其他估计数一样，由于X低效率而产生的福利损失是最重要的。

6.3 地方政府服务

迄今我们已检查了政府特许的和政府管制的垄断企业的X效率小于其面对产品市场竞争的同行的证据，前者如那些能够轻易地把较高的(燃料)成本转嫁给客户的公用事业。两种企业的情况表明，政府的力量造成企业感觉到追求X效率的压力减轻。政府影响企业X效率的另一种方式是通过各级政府间的援助，那就是一级政府对另一级政府的财政资助计划。例如联邦政府援助州政府和地方政府，为它们提供教育、公路、污水处理厂或职业训练项目的资金。像这样的拨款起什么作用呢？这种援助能增加州和地方政府创办的服务事业吗？它仅仅是用"联邦"的钱代替"州"或"地方"的钱？或者它鼓励州和地方政府浪费别人的钱？

6.3.1 西尔克曼和扬

理查德·西尔克曼和丹尼斯·扬(Richard Silkman and Dennis Young，1982年)研究地方政府是否受"他人金钱影响"的问题。这两位作者研究了1317个学校地区提供当地学校交通运送服务以及749个地方公立图书馆提供图书服务的成本如何受到由地方财政收入支付占各级政府间资助（来自州或联邦政府)的百分比的影响。如果政府间的资助提供了地方政府的收入，他人金钱的影响可以预言会使地方政府争取成本具有(X)效率的刺激减小。在此种环境中，地方政府官员在支出时较少对当地的纳税人负责。当地政府官员就成为

“免费搭车者”。也就是说，他们能够享受向他们选民提供商品与服务带来的好处，而不用从地方税收中出钱购买。

在答复地方政府是否受到他人金钱影响的问题时，两位作者首先估计在最有效的和一般的学校地区和图书馆中提供每一个服务单位的成本。估计“成本边界”和从边界的偏离程度揭示，最有效率的供应者提供学校交通运送服务的成本每个学生每年为67.07美元，而一般学校地区为142.72美元。最有效率的地方图书馆提供每册图书-小时的成本为0.24美元，而一般图书馆的成本为0.78美元。两位作者然后使用需求和供给条件以解释成本偏离“边界”的原因。就学校交通运送服务而言，需求条件包括服务的税收价格、收入以及就学年龄占人口的百分比。供给条件包括接收特殊教育（残废学生）在校人数的百分比，和外包给私营公司的学校交通运送服务的比例。对于图书馆服务，需求条件包括服务的税收价格、收入和财产价值。供给条件包括人口密度和图书馆分支机构的个数。此外，使用地方收入中提供服务的份额以测算对X低效率的刺激。两位作者报告说，当学校交通运送服务的成本由地方负担的份额上升10%时，每人偏离成本边界的程度减小2.8%或0.74美元。就地方图书馆来说，偏离边界的程度减小3.4%。一个有二万人口的一般学校地区当地方收入负担的份额增加10%时将发现节省经营成本大约4.5%。地方政府看来受到他人金钱的影响。也就是说，当那些支出的公共基金在更大程度上由他们的行为承担责任时，促使人们追求X效率的动机增加了。

6.4 交响乐团

正如联邦政府及其在各州的分支机构对地方政府给予政府间援助那样，它们也提供基金资助许多私人团体的活动，包括艺术表演团体。这些团体似乎特别易受市场力量的伤害，因为虽然它们为其投入支付市场价格，它们常常发觉难于（如果不是不愿的话）增加这些投入的生产率。它们的服务性质常常是劳动密集的，而难以变成资本密集的。政府和私人津贴就这样用来弥补“收入的缺口”，也就是弥补成本与收入之间的差距。

6.4.1 格罗布曼和布克

当然，政府有责任适当地提供财政资助。（但本书的准备和写作并没有得到政府的任何直接资助！）如果做得不适当，X 低效率行为不加处罚就将不必要地增加纳税人的负担。格罗布曼和布克（Globerman and Book，1974年）使用1971—1972年加拿大 33 家交响乐团和 23 家剧团为样本，调查加拿大政府的津贴是否影响 X 低效率。因变量是每次演出的成本。因变量包括演出次数、列为“重大”演出的演出百分比、观众人数、演出季节长度、投入价格和测算政府津贴效应的变量。作者选择用艺术表演团体的年龄作为测算最后一点的手段。这个测算办法的基本理由是历史较久的团体在如何填写拨款申请方面有更丰富的知识，政府也更愿意给予它们支持，因为已经对它们的艺术活动“投入”了基金。换言之，人们相信

历史较久的团体被认为较少为它们的成本负责。

对交响乐团和剧团用普通最小二乘法测算表明，团体年龄与每次表演的成本存在具有统计显著性的相关性。这两位作家的结论是，如果政府更密切地监督历史较长的艺术表演团体，它们的X低效率就会减小。

6.5 航空公司

6.5.1 普斯特

普斯特(Pustay，1978年)测量了11家航空公司在1965—1974年间每年的相对航运表现。普斯特相信，航空行业对于研究管制和企业表现是一个很好的对象。第一，在民用航空委员会管制下，市场价格和以全行业平均成本为依据的全行业利润率都固定下来。这样，管制向每家航空公司提供相同的刺激。第二，一旦你解释了航线结构的不同，而它们的投入与产出基本上是同质的。

但是，虽然管制可能向每家航空公司提供同样的"胡萝卜"，它也使航空公司隔绝竞争的"大棒"。管制可能就这样引起X低效率，因为没有适当的内部(胡萝卜)和(或)外部(大棒)的刺激。在一篇一般讨论管制者的行为，并让人联想起戈洛普和卡尔松以及乔斯科论文的陈述中，普斯特评论道，管制者监督成本，目的只在于确信这些成本的发生是出于"诚实、有效和节约的管理——一句话，他们的监督主要从法律上着眼，而不是从经济上着眼。"他的结论是，X低效率的可能性是存在的。他的研究试图测量航空行业中各企业在

1965—1974年期间每年X(低)效率的相对数量。

普斯特根据其成本对行业平均成本的比率，测量了任何航空公司任何一年的X低效率程度，只要“平均”的航空公司与在调查的航空公司飞行同一类型的航线。耗用的成本有三类：飞行、资本和乘客成本。飞行成本包括飞机运行、直接维修和保险的支出。在资本成本中飞机成本占了85%到90%；资本使用的效率部分依赖于停飞时间与飞行时间的比率。乘客成本是按每单位乘客英里收入的一般服务和管理费用来衡量的。

对这三类成本的每一类来说，效率指数都是指平均行业成本与个别厂商成本的比率(航线结构相同)。然后，通过对三类成本中的每一类按其在航空公司经营中的相对重要性进行加权，可以算出每家航空公司每年的综合效率指数；权数反映了整个1965—1974年期间每类成本的“平均”重要性。

对于整个10年期来说，普斯特的数据显示，如果平均水平的航空公司与效率最高的航空公司具有相等的X效率，那么成本可能节省12%到15%。这项研究清楚地显示了效率指数的“相对”性质。就是说，把X低效率视为“最高”效率厂商和平均效率厂商的成本之差意味着，“最高”一词是在绝对意义上使用的。因此，如果“最高”效率厂商的生产成本高于某一技术最低水平，那么12%到15%这个数字显然低估了节约成本的潜力。另外，12%到15%的成本节约是大还是小呢？就是说，它是否证明管制会在一个行业的厂商之间产生生产成本上的较大差异，或者证明管制减少了这种差异？正如普斯特所指出，对后一问题的回答要求我们检验非管制厂商间的成本差异。

6.6 结论

本章考察的证据充分表明，政府管制对X低效率的影响看来是巨大和多方面的。在下一章中我们将考察所有权形式的影响。

7 经验证据：所有权形式

7.1 导言

传统微观经济理论假定厂商总试图使利润极大化，即厂商一心一意追求利润。相应地，该理论隐含地假定企业中的每个人都具有或会被促成这种倾向。这些假定部分源于对厂商的一种观点：所有者占据着"金字塔"的顶端；而后这些所有者选择能反映并执行其追求利润极大化意愿的董事会；董事会转而雇佣经理，后者间接通过前者的指令按所有者的利益行事(董事会将解雇不能照此行事的经理)；经理们又发布命令给下属，后者又依次发布命令给其下属，以确保所有者的意愿得以实现。

事实上美国的许多公司是"经理控制"的，而拥有公司的形形色色的大量股东，其利益并不具有指导这些公司经营方向的支配力量。换言之，经理们对下列问题有一定的相机抉择权：他们将付出多少体力和脑力自觉地服务于公司利润极大化的目标，又将付出多少精力以实现其他明确的——对他们自己或对别人来说——目标，以及仅仅由于漠不关心会造

成多少浪费等。

微观经济理论基本上是一个关于市场经济和私有企业——它们构成私有市场——的理论。公有企业如何与私有企业相比较？公有制本身减少X低效率吗？经理控制本身减少X低效率吗？本章将考察有关这两个问题的经验证据。

7.2 所有者控制与经理控制的企业

7.2.1 谢尔顿

我们很少有机会研究这样一群企业，它们具有相同的产出——生产相同类型和相同质量的产品——并使用相同的资本和非资本投入。然而1967年约翰·谢尔顿（John Shelton)就进行了这样一项研究。他研究了一群企业的经营状况，它们在上述这些基本要素方面完全一致，唯一的区别是所有权形式不同。谢尔顿的研究对象是一个大型的全国性快餐饭店联号，它们中的大多数是由一个在法律和报酬上都独立于母公司的特许经营人所拥有。这种饭店有两类：较大的一类是家餐式饭店，较小的一类是快餐店。无论是哪种类型，也无论饭店的所有权形式如何，母公司对菜单、烹饪法、配料、桌椅摆设、服务和结帐手续等方面都有详细规定。整个过程达到如此彻底的标准化，并受到如此严格的监督，以至于被选中的特许经营者都是一些以前从无饭店管理经验的人，因为据认为以前有过该业务经验的人会试图按自己的经营方式行事。有时，一家饭店的所有权会从特许经营者转入母公司手中，或者相反。当母公司亲自管理饭店时，它会从

自己的工作人员中委派一名经理，后者十分熟悉这类饭店的经营，并富有管理经验。

谢尔顿的抽样包括22家这样的饭店，其中的15家在他的研究阶段中经历了所有权形式的变化，即从特许经营人所有(FO)转成母公司管理(CM)①，或者相反。此外，6家饭店经历了两次所有权形式的变化，另1家饭店则发生了五次这种变化。因此，这项研究共包括53个"饭店单元"的数据。

在所有53个样本中，FO所有权形式的平均时间长度为38周，而CM形式下的平均经营时间为33周。在CM形式取代FO形式以后，销售额下降了7.3%。反之，当CM形式转为FO形式时，销量上升了19.1%。利润率方面的对比则更加惊人，FO形式下的利润率(利润/销售额)是9.5%，而在CM形式下的利润率只有1.9%。每周平均利润在FO形式下为271.83美元，在CM形式下仅为56.81美元，两者相差478%。此外，29个FO单元中仅有2个出现亏损，而24个CM单元中有11个出现亏损。而且，在CM取代FO形式的22种情况中，有18个出现利润下跌状况，2个利润增加，2个维持原状。另一方面，在FO取代CM形式的情况中，利润上升的有7个，下跌的有2个。在较小的快餐店单元中，FO形式下的每周平均销售额约高出19%。FO形式下的利润率约为16.2%，而在CM形式下约为6.4%。

一位母公司的经理人员对上述现象是这样解释的："……特许经营所有者对业务的细枝末节更关注，他们能更好地使用厨师和招待，他们减少了浪费。"尽管母公司任命的经理处

① FO、CM分别是franchisee-owner(特许经营人所有)和company management(母公司管理)的英文缩写。——译者

于严格的监督之下，并具有赢利动机（其薪水中的利润分成部分可高达33%，平均能达到15%），但独立的特许经营业主的经营绩效却更为出色。谢尔顿对此总结说，用你自己的钱投资而且仅从利润中获取报酬这点造成了在减少X低效率方面的差异。

7.2.2 蒙森、朱和库利

利用行业资料对所有权形式-利润关系进行的一项研究是由蒙森、朱和库利(Monsen,Chiu and Cooley, 1968 年)进行的。这些研究者从以下12个制造行业中各自选了 6 家企业：肉类产品、罐头食品及储藏、化工产品、医药、炼油、钢铁、有色金属、电气、汽车、飞机、工业机械和商用机械。所有这72家企业根据其 1963 年的销售额都被载入 1964 年 7 月《幸福》杂志的 500 家美国最大企业的名单中。这 72 家企业的资料涉及 1952—1963 年的整个时期。

在每个行业内，一半（3 家）企业是"所有者控制"，另一半是"经理控制"。所有者控制的企业是指业主积极参与企业经营并至少拥有10%具有投票权的股票，或者业主并不积极参与经营管理但却至少拥有20%投票股的企业。经理控制的企业是指没有一个所有人拥有 5%以上的投票股，并且近来也无业主控制迹象出现的企业。

为了估计利润-所有权形式的关系，这些学者拟定了行业、企业规模和年份三个变量。数据取自穆迪的《工业手册》和那些公司的年度报告。每种所有权形式的观察样本总数为每年 36 个（12个行业，每个行业 3 家企业），12年总计432个

样本。然后通过方差分析和三向协方差分析的平衡固定模型来进行估计。

利润或业主投资的收益，是以净收入占净资产的比值来计量的。所有36家所有者控制的企业在整个11年期间的利润率均值为 12.8%，而 36 家经理控制企业的平均值是 7.3%。此外，在所有 12 个行业中，所有者控制的企业都要比经理控制的企业经营得好。学者们认为，最好的解释是“存在两种大相径庭的激励机制——一是为所有者，而另一个是为经理”。这里的一层含义是，利润极大化的假定对于经理控制的企业并不十分适用。关于X低效率几位学者陈述说：“……莱宾斯坦报道的证据表明，激励因素对整体经济效率的影响比传统微观经济理论对配置效率的关注引导我们所相信的要重要得多。因此，所有权与管理的现代分离可能会对企业经营的动力产生影响——也许会比配置效率产生更大的后果……”

7.2.3 麦凯克恩

对所有权形式的一种略为不同的分类方式是把企业分成“所有者管理”、“经理控制”或“外部控制”三类，并同时从利润和增长两方面来衡量企业的经营状况。这正是威廉·麦凯克恩(William McEachern,1978年)所做的，他从1964—1973年期间的医药、化学制品和石油精炼三种行业中各选了16家企业作为样本。样本包含 8 家所有者管理的企业、26家经理控制的企业和 14 家外部控制的企业。在麦凯克恩定义的所有者管理的企业中，主要股东至少拥有 4% 的投票股，同时又是经理。而经理控制的企业则定义为没有一个主要股东拥有4%

以上的股票。外部控制企业的主要股东则不是经理。麦凯克恩因而在其模型中假定，一个主要股东的动力和表现将因他(或她)是否是经理而不同。

麦凯克恩的假说认为，外部控制企业中的主要股东无论同兼任经理的所有人相比，还是同经理控制企业中的经理相比，他从企业增长中获得效用的可能性都要小。据认为在后两者中，这些股东的自身利益更可能与企业的发展规模相联系。因此，外部控制企业积极追求增长的可能性较小。

通过运用普通最小二乘法，并拟定企业年龄和其行业属性两个变量，他的研究结果显示，所有者管理和经理控制的企业的增长率都比外部控制的企业高。然而，该结果只对所有者管理的企业存在统计上的显著性。换言之，按增长率高低三类企业的顺序是：(1)所有者管理的企业，(2)经理控制的企业，(3)外部控制的企业。同时他还指出，医药行业的增长率要明显高于化学制品行业，而后者与石油行业的增长率却无显著差别。最后，他指出增长率随企业年龄的上升而下降。

不过，增长率并不必然意味着企业在更多地赢利。我们不妨回忆一下在第3章的马里斯模型中所进行的讨论：增长在超越了一定临界点之后会降低资本收益和团体的生产率从而降低利润。然而麦凯克恩的数据表明，具有较高增长率(所有者控制)的企业其投资效益并不差。1963—1972年的数据表明，所有者管理的企业在该期间所持有股票的收益率比经理控制的企业高（即其主要股东集团的收益率比经理控制企业的高）。麦凯克恩指出，这意味着如果所有者管理的企业对增长的追求超出了使利润现值极大化的界限，那么他们显然已经

不仅仅停留在通过以其他方式提高效率来弥补利润上的损失这一步了；"也许X低效率并不像莱宾斯坦所讨论的那么多。"

7.2.4 纽豪斯

医疗保健的提供也可以在不同的所有权形式下进行。例如，一些医生是单独开业者，另一些则在大型医疗门诊所工作。由于单独开业的医师有权自由分配所有的税后收入，很明显他有工作更长时间和控制成本的动机。然而，在大型医疗门诊所工作的医生却很可能面临成本和收入分享计划。在这种情况下他既不承受全部费用的负担，也不享有自己工作努力的全部收益。假使其他条件不变，他的行为就很可能像一个"搭便车者"。另外，大型医疗门诊所也更可能获得政府补助以弥补成本，从而使这些成本相对高昂的诊所免于因市场机制的作用而被剥夺生存之地。因此，与单独开业医师相比，有两方面的原因使我们确信大型医疗门诊所会遭受更多X低效率的危害。首先，在一个大型医疗门诊所工作的单个医生较少具有使成本降到最低和付出最大工作努力的动机；其次，效率差的门诊所被排斥出市场的可能性降低了。正如普里莫克斯的研究所显示的那样，大型医疗门诊所据认为能得益于规模经济，但却会受累于X低效率。

约瑟夫·纽豪斯(Joseph Newhouse, 1973年)对这些问题作了研究，为此他使用了20位单独开业医师或独立的专业团体(私人开业)、以及3家大型医疗门诊所1969年的数据(最大的独立专业团体由5人组成)。私人医师的数据资料来自汇总的调查表，而门诊所的有关数据则是纽豪斯本人估算的。

虽然数据的确表明门诊所比不实行成本共享的医师招致更高的成本，但对此经常给予的一种解释是，门诊所的较高成本来自于它们的教学计划。该论点的推断如下：(1)门诊所和私人开业的唯一区别(一种假设)就是存在一项教学计划；(2)经营门诊所的成本更高，因此推论便是(3)实施教学计划的成本很高。然而，纽豪斯指出门诊所的高成本与其教学计划无关，他将这种高成本归究于门诊所中存在的较大程度的X低效率。

例如他报告说，在医疗门诊所每次就诊的管理费用平均为14.24美元，而在私人医生处为4.54美元。这些管理费用包括开帐单、门诊记录、办公和护理的成本。其次，通过使用普通最小二乘法，并拟定门诊次数和门诊次数的平方为变量，纽豪斯发现成本共享计划使每次门诊的薪水成本提高了2.55美元。在此基础上他指出，一个每月进行400次门诊的单独开业医师，其每次门诊的薪水成本（2.48美元）比实施成本共享计划团体的成本低，不管后者每月所做的门诊次数是多少。（例如，1000次门诊的每次薪水成本是2.85美元。）换句话说，规模经济效益被很快耗尽了，因而大型医疗门诊所的X低效率的成本经常超过其规模经济的收益。（这一结果类似于普里莫克斯的发现，即X效率带来的赢利超过了规模经济的效益。）

运用同样的模型和同样的普通最小二乘法步骤，他又估算了成本共享计划对薪水成本主要成份（医疗记录、预约成本、开帐单和租金）的影响。结果是每次门诊记录费增加了0.24美元，每次预约门诊成本增加了0.23美元，以及每次门诊开帐单和租金成本分别上升了1.11美元和0.96美元。第

三，1967年4月单独开业医师平均工作了218小时，而在大型门诊所工作者的平均工作时间为197小时。(这种情况是配置低效率还是X低效率呢?)

纽豪斯也提出了高成本是否反映高质量的医疗水准这一重要问题。虽然他没有给出一个结论性的答案，但他确实指出62%的更高成本来自诊所内的医务辅助活动，如获取医疗记录、开制帐单和订预约等。因此他推断出，较高的成本反映了X低效率而基本上并不代表更高的医疗质量。

7.2.5 戴维斯

另一个对医疗保健领域进行过研究的是卡伦·戴维斯(Karen Davis, 1973年)。戴维斯检验了三种关于医疗费用上涨的理论：需求拉上、劳动力成本推进和成本加成补偿理论。在她的成本加成补偿模型中，X低效率在成本函数中成为一个"移动"参数。也就是说，平均成本会受到被置于成本补偿计划下的病人数量(qm)、不被置于该计划下的病人数量(qn)以及X低效率的影响，因此X低效率能使整个成本曲线向上移动。

但是，X低效率既能使成本曲线上移，同时又能增加医院的利润。这种情况在病人总数中被置于成本补偿计划下的人数比例超过成本补偿比例的倒数时就会发生。她对此解释说，假定k是成本补偿的比例，因此如果成本加成因素是5%——对每1美元的成本医院将获得补偿1.05美元——那么k=1.05。因此，如果qm/(qn+qm)>1/k，X低效率就会同时增加成本和利润。我们不妨看一个戴维斯所举的例子：

如果成本加成因素是5%，那么要使X低效率能增加赢利就得使95%以上的病人置于成本补偿计划之下。图7－1中的曲线Ⅰ是使X低效率既不赢利也不招致亏损的k和置于成本补偿计划下的病人人数百分比的所有组合。所有在曲线上方的组合反映了X低效率是一种对医院颇有吸引力的选择，而在曲线下方的点表示X低效率无利可图。

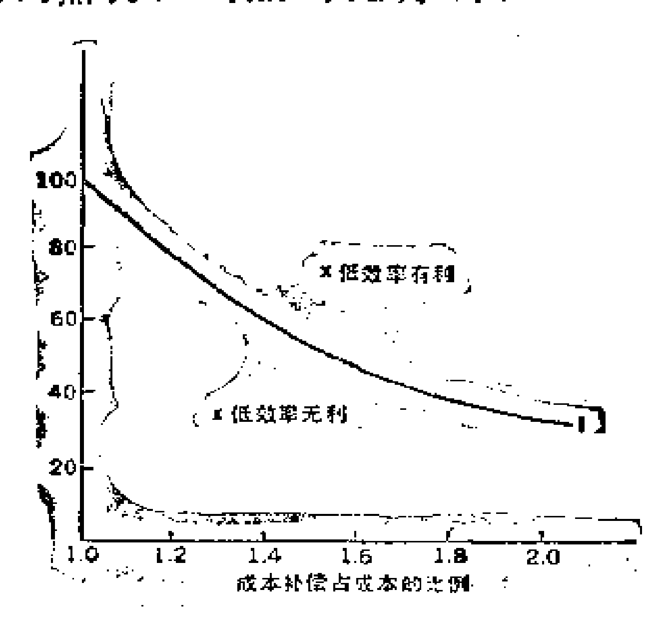


图7-1　成本补偿计划和X低效率的吸引力

戴维斯运用普通最小二乘法及1965年、1967年和1968年的数据估测了所有美国各地医院中每个病人的医疗成本。成本补偿计划的效果是用置于该计划下的医院费用比例来衡量的。此外，她还使用1967年和1968年两个虚拟变量。(老年医疗保健制度是1966年开始实行的。) 方程中的其他变量则用来测量医院床位的使用率、医院的平均工资率、医院的规模(床位数目)和平均住院时间等。回归分析中包括了测量成本补偿比例的变量，但没有包括1967年和1968年两个虚拟变量；该分析是分别以三年中的每一年为时期单位，并且

以整个三年为一时期单位来进行的。只有在汇集了三年所有数据的回归分析中成本补偿变量才呈现出统计上的显著性(而且为正)。当对所收集数据进行的回归分析中也包含时间虚拟变量时，成本补偿变量虽为正，但却并不显著。不过，两个时间虚拟变量都具有统计显著性，而且都是正的。

成本补偿变量的统计显著性可以解释为，当置于成本补偿计划下的病人人数百分比上升时，成本也相应增加。时间虚拟变量的统计显著性则表示老年医疗保健制度对更高成本的一时性影响无法被排除。虽然更高的成本也可因技术变化和(或)产品质量的改进引起，我们仍然想强调对成本补偿和时间虚拟变量的计量结果与X效率理论是一致的。最后，戴维斯的理论和统计分析结果指出，X低效率可以是"自愿性的"，也可以是"非自愿性的"。

7.2.6 居南卡

提高农民的经营状况(尤其是在欠发达国家)引起了两种不同的反应和对策建议：(1)土地实行社会化，和(2)限定拥有土地的最高数量和将所有权转移给佃农。

一方面，有时人们认为大农场更倾向于追求利润极大化。另外，大农场的所有者能获得更好的信息，因而更可能利用最新技术；他们也更容易进入不完全和未开发的资本市场。另一方面，限定农场最大规模建议的理由是，在超过某些界点以后，增加农场的规模既不能提高生产率也不能提高别的经营标准。将所有权转移给佃农的观点则认为，拥有自己耕作的土地会提高农民的经营绩效。如果不能获取自己劳动的

全部收益，以及对土地的未来使用更加没有把握的话，佃农很可能会减少高生产率的资本投入，而进行那些需要更长回收期的投资。

居南卡(Junankar,1976年)以1968—1970年间印度旁遮普邦的农场为对象考察了农场规模和所有权形式对X效率的影响问题。所有权形式是以出租土地占土地总面积的百分比来衡量的。因此，一个农民拥有其耕种的土地越多，就代表他的所有权也越多。X低效率则以主要作物小麦（按所有者和佃农既定的投入量生产的）价值的差异来衡量。于是，居南卡通过比较农场中使用类似数量的劳动力(小时)、肥料、拖拉机、耕牛获得的产出来衡量所有权的影响。他的统计分析针对25英亩以下的小农场和25英亩以上的大农场两种类型进行。他的研究结果表明，农场规模本身并不影响X效率，即就每单位投入的产出而言大农场并不比小农场更多。然而，在大农场中，拥有较多所有权会提高X效率，即在大农场中拥有较多所有权会增加既定投入量的产出。

7.2.7 蒂默

行业生产边界和X低效率

各种研究通过估测一种行业的生产边界来测量X(低)效率和技术(低)效率。估测生产边界是通过"设置"一个生产函数来进行的。例如，可以建立下列形式的科布—道格拉斯函数：$\ln Q = \ln a_1 + a_2 \ln K + a_3 \ln L + e$，其中$a_1$是常数，K和L分别代表资本和劳动，e是误差项，$a_2$和$a_3$则分别是资本和劳动的系数或产出弹性。因为边界生产函数是最大（产出/投入）

份额的点的轨迹，因此该函数受误差项必须是非正的约束。就是说，由于误差项是实际产出(Q)与可能或预计的产出(Q^*)之间的差异——$\varepsilon = Q - Q^*$——因此误差项必须限定是非正的。换一种说法便是，生产数的建立将受到如下约束，即可能的产出必须不低于实际产出。我们可以用一个线性规划来处理该函数。

一旦估测出该生产函数的系数，在给定每个企业投入的情况下，企业可能的产出就可以计算出。于是，每个企业的X低效率(有时被称作技术低效率)水平便可用Q/Q^*这一比值表示，因为它反映了每单位的投入还能获得多少更多的产出。(在本章的后面部分我将讨论这种衡量X低效率方法的缺陷，以及X低效率和技术低效率间的差别。)

蒂默对美国农业的研究

彼得·蒂默(Peter Timmer,1971年)曾对此作了一项研究。蒂默估计了美国州一级农业的生产边界。其中，产出是以1960—1967年间农业产量的美元价值来衡量；资本是用经济价格指数折算的当前农业营业费用来测量；土地是以其卖价来测算，因此这就假定土地生产率的差异反映在其市场价格之中；劳动力包括所有的农场员工，即农场主家庭和雇用的工人。

蒂默为了减少偏差和站在局外人的客观立场上测量效率，他从每个州中剔除了2%效率最高的企业，由此得出的结果是：南达科他州的效率(Q/Q^*)等级最高，为99.1%；西弗吉尼亚州最低，为81.0%；平均效率等级为93.4%。也就是说，每个州的效率平均低于生产边界6.6%。

对测量方法问题进行调整后，该数字应降低到3—4%。当然，综合各州的数据并不能使我们从中观察出州内部各企业间在技术或X效率上的差异。蒂默认识到这种差异存在的可能性以及他的数据对于反映州内部X低效率的无能为力。

然后，蒂默把这些效率等级作为根据几个自变量作回归分析的应变量来预测效率。其中一个自变量是佃农占农民总数的百分比。回归分析的结果显示，当佃农的百分比上升时州的效率等级也会上升。蒂默评论说，这可能是"因为非南方各州的年轻农民为了积蓄足够财力购买自己的农场而付出了额外的努力和动力"。由此得出的研究结论是，佃农中的X效率更高是因为他们有动力，劳动也更卖力。虽然许多其他方面的因素也可能导致对生产边界的偏离，但蒂默却选择了X效率作为原因。

蒂默也用普通最小二乘法的回归分析技术估计了行业生产函数。这一技术提供了一种"平均"生产函数而不是边界生产函数。研究结果在某些有趣的方面与边界生产函数存在差异。首先，边界生产函数中的资本系数比平均生产函数中的系数高出30%。其次，边界生产函数中的劳动系数比平均生产函数中的系数提高40%至70%。蒂默对此的解释是，最好的农场使用较少的劳动力，因此其劳动的边际生产率较高。而这反过来又是良好的管理和采用资本密集度较高的技术造成的。

这些研究结果似乎与居南卡的不一致。但是，在资本市场高度发达，地主和佃农拥有同样多的进入这些市场的权利的美国，升入较高的社会阶层是较为普遍的现象，因此佃农积蓄财力以购买自己农场的动力更大。而在印度这个比较传统

和资本市场欠发达的国度中，佃农就较少拥有这种机会。另外，居南卡还在其文章中提到，为了阻止佃农获得土地所有权的可能性，他们被不时地迁移着。因此，印度的佃农不愿像地主一样付出那么多的努力是可以理解的，正如作为蒂默研究样本的美国佃农愿意比地主付出更多的努力也是可以理解的。因此，美国和印度佃农获得机会的不同可能性促成了有利于前者的不同激励机制。

7.2.8 夏皮罗和马勒

夏皮罗和马勒(Shapiro and Muller, 1977 年)用类似于蒂默的方法估计了一群坦桑尼亚农民的X(低)效率和“现代化”的关系。(他们交替使用X(低)效率和技术(低)效率这两个词。)虽然他们在研究中没有对具有不同所有权形式(或不同市场结构)的农场的这种关系进行估计，但与蒂默所用方法的相似性促使我们在此讨论这篇有趣的论文。

夏皮罗和马勒的数据取自1970—1971年间坦桑尼亚一个55平方英里行政区中的 26 个棉花种植农场(占人口的10%)。关于劳动(工作小时)、土地(种植棉花的英亩数)和产出(以坦桑尼亚货币计算的棉花销售额)的数据是在一年中每周收集两次。这里没有包括资本数据，因为(1)资本很少使用，而且(2)多数农民对化肥和杀虫剂的使用不当。两位作者感到对以上数据作标准化处理所需进行的纠正过于复杂，而且显然得不偿失。(否则他们是会这么做的。)

现代化方面的数据是使用调查表和采访的方式来收集的。通过此法所获得的关于85个变量的数据反映了农民们采

用可能得到的现代经营管理手段的相对程度。他们采用格特曼量表①技术，得出了总共含45个项目的9个现代化量表。每个农民在每个量表中的现代化分数是以他在该大类中采用的项目数量计算的。这9个量表分别测量下列内容：(1)棉花生长注意事项方面的知识，(2)投入和产出价格方面的知识，(3)对当地农业官员的了解，(4)对农业信息的搜寻，(5)一般作物的生长，(6)农业投入物的使用，(7)农业财产，(8)家用器具设备，以及(9)房屋及其组成物中的固定部分。

两位作者估测了一个仅含劳动和土地两个自变量的（科布—道格拉斯）边界生产函数。该函数导出了每个农场的可能产出(Q^*)，并因此提供了估计所有农场经营中的X低效率程度(Q/Q^*)的基础。然后，两人揭示了X效率与一般或总体现代化程度的关系，以及与9个格特曼量表指标中的每一个的关系。X效率与前者和后者都正相关，其中与总体尺度的相关度为0.566，统计显著水平为1%。其他显著性在1%水平上的量表分别是量表2、3、8和9。显著水平为10%的为第1、5、6和7号量表。（量表4的显著水平为20%。）

他们估测的边界生产函数导出了下列（对数）方程：$Q=5.343+0.050K+0.803L$。用普通最小二乘法对该（对数）方程进行估计，得到一个平均生产函数（“平均”企业的生产函数）$Q=5.470+0.291K+0.690L$。换言之，通过对这两个生产函数结果的比较我们可以得出下列推断：最好的企业比平均企业具有更高的X效率，这种状况是通过更高的劳动生产率取得的，而不是通过更高的资本生产率或通过生产函数

① 格特曼(Guttman,1916—)，美国心理学家，格特曼量表的发明者，该量表是作为心理学、社会学中量度心理态度和性质之用的。——译者

的"中性"位移取得的。(中性位移可以从常数中"捕捉到"。因为两个方程中的常数基本上是一致的，所以结论是中性位移并没有发生。)

两人随后报告了他们估计一个包含现代化变量的平均生产函数的结果。现代化变量既可以常数出现，也可作为与劳动和土地发生交互作用的因子出现。当它以常数出现时，现代化被假定会影响生产函数的截距，而不影响任何要素的生产率。因此截距代表了现代化的"全面"影响。也就是说，现代化产生了在某种程度上提高农场产出率的外部因素。另一方面，当与某一投入要素发生交互作用时，现代化程度的改变被假定会提高既定数量的那种投入要素的边际生产率。

八个不同回归的结果被显示出来，每一个代表对现代化变量的不同使用(组合)。总的说来，其结果可以概括如下：首先，土地和劳动的边际生产率都是正的，且基本上具有10%的显著性。其次，当以影响截距的常数出现时，现代化一般会显示出一个正的全面影响，并具有0.10的统计显著性。第三，现代化程度的提高一般会减少土地的边际生产率。第四，现代化程度的提高一般会增加劳动的边际生产率。换言之现代化——代表一种对高效率的愿望，而不仅仅是知识的存在——基本上是"劳动推进型"的，它是解释这些农场样本中生产率差异的重要决定因素。

7.2.9 技术和X(低)效率的比较①

蒂默、夏皮罗、马勒，以及其他一些人都交替使用技术

① 关于X和技术效率概念的异同，参见莱宾斯坦(1977年)的论文。

效率(TE)和X效率(XE)两个词。它们虽然相似，但并不完全一致，而正因为它们有差异，对其中一个的测算并不必然意味着对另一个的测算。首先，TE是一个源于新古典理论的概念，而XE在若干基本方面偏离了新古典理论。第二，TE意指极大化行为，而XE则允许同时存在极大化和非极大化的行为。第三，TE把厂商视为行为的基本单位，XE则把个人视为行为的基本单位。第四，TE假定厂商是个统一的决策单位，具有管理人员所制定的明确目标；XE则允许厂商具有几个或许是冲突的目标，而且这些目标可以不被明确表达出来。

最后，TE是由诸如管理者和工人之间的不同技能等技术因素造成的；或者，由于使用的是一种特殊的技术，虽然该技术本身已被充分利用，但它的使用效率仍会低于同行业中其他企业对类似技术的使用。这里的问题在于，不同年代的资本在任何时候都可能被使用，这不是因为经理们忽视了这一问题，而是因为技术的扩散不是瞬间完成的，因为不同企业对所需资本的可得性不同，或者因为资本的生产率达到其最大潜力需要一定时间。因此，我们会看到这样的例子，被体现的更快的技术进步会以更大程度的技术低效率的面貌出现。然而，X低效率却不是因技能的参差不齐，或因技术的扩散需要时间所致；而是因为对技能的使用不同，因为工作规范可以不同，因为工作的努力程度会有某种相机抉择要素，也就是因为对成本和生产率的重视程度并不总是达到可能达到的地步。因此，由行业生产边界测量的技术低效率的存在并不必然意味着X低效率的存在。

因此，每项研究必须作为一个独立的单元进行，以分辨

所研究的到底是XE还是TE。例如，夏皮罗和马勒指出，农场经营绩效的差异源于是否愿意采用可能采用的有关生产的现代化手段，或是否愿意，用作者的话来说，"达到技术高效率"。因为这些手段都是可能获得的，因此似乎应把这视为一种X效率的情况。我们不妨再回想一下9个量表中关于投入和产出价格的知识的那个量表。缺乏这些价格方面的知识是否会导致配置低效率呢？例如，如果厂商并不了解投入要素的相对价格，他们就不可能运用使成本极小化的资本/劳动比率。也就是说，即使厂商按其等产量线生产，他们也不可能在该等产量线上 $MP_l/MP_k=P_l/P_k$ 的那一点进行生产。在这一例子中，厂商虽然达到了X高效率，却同时表现出配置的低效率。然而，如果从他们缺乏关于投入要素价格的知识是由于他们不关心市场状况这一角度来看，我们会得出结论说，这是一种以配置低效率形式表现出来的X低效率。毫无疑问，这些问题是需要我们仔细分辨和处理的。

7.2.10 格雷戈里和詹姆斯

被体现的技术进步对估计X效率和技术效率的影响当然是一个重要问题。然而，这种被体现的技术进步已经不仅是指技术变革被体现在最新的技术中，也是指该技术已被充分使用。至少它意味着最新的资本必须与比伴随旧资本的更高劳动生产率相结合。这就是那些"年代"资本模型所作的假定。这些模型认为X低效率为零。就是说，劳动生产率以机械的方式形成，一旦资本存量的数量和年代已知，它就几乎同时被"预先决定"了。于是我们又回到了假定实际产量总是

最高产量并取决于劳动和资本投入要素以及技术的新古典生产函数。

格雷戈里和詹姆斯(Gregory and James,1973年)利用澳大利亚46种行业中的116家新工厂的样本检验了年代资本模型。这些工厂是在1953至1958年间开始经营的，格雷戈里和詹姆斯对年代资本模型的分析集中在1957—1958年间，但只包括在1956—1957年间开始经营的工厂。新工厂是指那些建于新址的工厂。数据的收集来源于一张调查表，该表的回收全凭工厂雇员的自愿。同夏皮罗和马勒论文中的情况一样，虽然格雷戈里和詹姆斯的论文并不是关于所有权形式的，但它却涉及我们正在讨论的题目之一：实物资本和X效率因素对生产率差异的相对影响。

对年代资本模型的检验是通过在每种行业内比较新工厂与原有工厂的劳动生产率数据来进行的。劳动生产率是以每个工人增加的价值来衡量的。由于工厂生产产品的多样性，我们无法对更好的衡量标准——每个工人的实物生产率——进行估计。

他们对年代资本模型进行的第一项检验是比较每个行业内所有新工厂的劳动生产率差异。劳动生产率的差异总是可以预料到的，因为存在资本/劳动比率的(合理)差异，这是由例如对未来价格的不同(合理)预期等因素造成的，或因为行业中存在错误的规范，也可能是由于不同工厂内具有不同的产品组合。然而，在考虑到这些因素以后，年代资本模型仍可得出这样的预测：与所有工厂间的劳动生产率差异相比，所有新工厂间的劳动生产率差异相对较小。

这种比较可以在25种行业中进行。在半数以上的这些行

业中，生产率最高的新工厂的劳动生产率是最低的新工厂中劳动生产率的三倍以上。虽然无法进行与老工厂的劳动生产率差异的比较，两位作者断定两者间的变动幅度很可能会比我们从年代模型中预计的大。如果我们考虑到因为只有新工厂在进行比较，所以就工厂的年代而言劳动生产率的差异已被"标准化"了，那么我们对这一结果的解释能更加清楚。因此，按照作者的观点，新工厂间的劳动生产率差异反映了一些非年代因素，包括X效率的影响。

他们进行的第二项检验是比较新工厂与行业的平均劳动生产率。根据年代模型的预测，新工厂的劳动生产率应超过行业的平均水平。而且，资本年代因素比非年代因素对劳动生产率的影响越重要，新工厂的劳动生产率与行业平均水平的差异也应该越大。于是，两位作者将新工厂的生产率水平除以行业的平均水平。结果是，在46种行业中，24种行业的新工厂生产率超过行业平均水平，而在其他22种行业中情况正相反。另外，在116家工厂中，59家工厂的劳动生产率超过行业平均水平。因此在约半数的情况中，新工厂的生产率超过行业平均水平。所有46种行业的综合比率(新工厂的劳动生产率/行业的平均劳动生产率)是1.103，这说明年代因素是起作用的。

两人还检验了另一种可能性，即在选出的几种较大行业中的较强年代效应会使整个经济的年代效应增大（超过上述1.103的水准）。他们的检验是通过对该比率（新工厂的劳动生产率/行业的平均劳动生产率）作加权平均处理来进行的，其中权数反映行业的相对规模(雇佣人数和产出量)。对于全部含46种行业的样本，该比率为1.15（按雇佣人数加权）和

1.18(按产出量加权)。由于规模经济因素在大行业中会更重要一些，该数字可能高估了年代资本的效应。然而撇开这一因素，索尔特(Salter，1966年)等年代模型的支持者都认为，一般可以发现新工厂的劳动生产率会高出50%至100%。

还有两个问题应该提一下。第一，新工厂的劳动生产率可能低于行业的平均水平，因为在达到最高效率以前需要一段学习和"安顿"的时间。两位作者对这一问题的检验是通过计算1956—1957年和1957—1958年间的一项比率（新工厂的生产率增长/行业的平均生产率增长)来完成的。如果学习和安顿因素对新工厂的相对生产率增长并不起重大作用，那么该比率应当随机分布在1.0左右。在可以计算该比率的 37 种行业中，18 种行业的这一比率低于 1.0。在 116 家工厂中，48%工厂的生产率增长低于行业平均水平。因此在这一样本中，学习和安顿问题看来并不是一个新工厂要花大力气来克服的障碍。

第二个要考虑的问题是，新老工厂间生产能力利用率的差异会使任何比较的结果难以解释。两位作者指出，新工厂的生产能力利用率约为80%。就当时而言，这一数字是具有代表性的，甚至会比预料的大。换言之，生产能力利用率的因素不会使研究结果发生偏差。因此，两位作者得出如下结论，新工厂是否体现了最好的技术和(或)其他诸如X效率等非年代因素对于决定工厂的生产率都起了同样的重要作用。

7.2.11 申

运用 1947—1959 年间马萨诸塞州制造工厂的数据，申

(Shen,1973年)指出，"经营最佳"工厂的Q/L和K/L两项指标都高于其他同等规模的工厂。然而，X效率对Q/L的提高更大于对K/L的影响，而后者的提高更多是投入要素替代的结果。申推断说，该结果是与技术扩散模型一致的，如果该模型包含X(低)效率因素的话。

7.2.12 布拉德利和盖尔布

同已介绍过的略有不同的一项研究是对西班牙北部巴斯克地区的私人合作社(称为蒙德拉贡)与同一地区的非合作社的比较。虽然布拉德利和盖尔布(Bradley and Gelb,1981年)没有发表对合作社和非合作社经营绩效的比较，但他们却引用了托马斯和洛根(Thomas and Logan, 1982年)的论文材料。从后者所显示的生产率、销售量增长以及利润等数据来衡量，合作社的经营绩效一般优于同一地区的非合作社。我只简单假定托马斯和洛根的结论，然后再介绍布拉德利和盖尔布的有关研究结果。

布拉德利和盖尔布是少数在研究中区分影响企业经营的"外部"因素和"内部"因素的人之一。在他们的研究中，外部因素指工资水平、工作安全感，内部因素包括企业的"巴斯克特征"，以及其自身的合作性质和工作条件。既是业主、又兼为雇员的蒙德拉贡合作社成员和比较传统的私人企业（管理团体）中雇员的一个区别是，只有15%的合作社成员将其工资收入视为优先或次优考虑的事情，而那种管理团体中的成员有37%。合作社成员多半将合作性质视为他们企业的最重要特征，而在管理团体中，工作安全感的重要性排第一位，

工作条件排第二位。由此可以预料，较之管理团体，合作社成员是更不愿意因工资收入而改变工作的。

布拉德利和盖尔布还指出，相对于管理团体而言，合作社成员感受到的劳资间鸿沟(辅助性的垂直关系)要小得多；他们很少支持工会；他们有很强的参与企业重要决策的意识；他们更强烈地感到，企业的成功依赖于他们每个雇员的工作努力程度，同事间的压力(水平关系)有利于激发努力的工作。布拉德利和盖尔布得出的结论是，合作社的环境提高了X效率(这可以用托马斯和洛根论文中所讨论的各种绩效标准来度量)。

7.3 公有和私有

7.3.1 引言

有一些理由使我们相信，公有企业的X效率会低于私有企业。首先，政府的征税权力可在需要时用来资助公有企业。第二，这些企业不能像私有企业那样自由分配企业的利润。也就是说，由希望获得利润促成的追求X效率的动力减小了。第三，公有企业常常是垄断者。然而，这里的证据却显示出复杂的结果，即在某个例子中我们发现公有企业具有更高的X效率，而在另二例中情况却相反。这种复杂的证据很类似于在论述关于公有权和私有权的广泛文献中所显示的情况①。这里所展示的三个例子是：(1)美国的自来水公司，

① 公有企业和私有企业经营绩效的差异概括在阿特金森和哈尔沃森(Atkinson and Halvorsen,1986年)、克鲁和克莱因多佛(Crew and Kleindorfer,1985年)以及德·阿莱西(De Alessi, 1980年)的论文中。

(2) 巴西的塑料和钢铁制造业，以及(3) 玻利维亚和印度尼西亚的采矿业。

7.3.2 布鲁金克

自来水公司，不管是私营的还是公营的，都受到利润率法则的管制。正如我们以前提到过的，这些管制至少会从两方面妨碍对成本的控制。首先，管制机构对利润的控制会减少所有者监督成本的动力。因为所有者不能全部获取利润，因而可以预料他们不会努力经营。其次，利润率管制的成本加成性质减少了经理人员追求成本极小化的动力。

私营或公营公用事业公司的经理们又都面临实现X效率的压力。一家私营公用事业公司的经理必须取得一定水准的X效率(利润)，否则就有丢掉工作的危险。公营公用事业公司的经理们也会面临同样的危险，如果其决策明显阻碍了对公众服务的话。然而，公营公用事业公司的所有者(纳税人)与私营部门所有者监督管理者的动机不同。一些研究者包括托马斯·布鲁金克(Thomas Bruggink)认为，公营公用事业公司的经理具有另一种追求X效率的动机，即在一个"不友好的经营环境"中，这些经理必须通过有效地经营公用事业来证明他们的管理才能和公有权的好处。

布鲁金克检验了这个观点，为此他采用了一个包含77家公营和9家私营自来水公司的1960年的样本。成本是以每生产一百万加仑自来水的营业成本来计量的。数据是由美国自来水厂协会提供的。为了检验所有权形式的影响，布鲁金克设定了产量、环境、分配、生产要素成本和管制因素等变量。

这些因素包括自来水产出量，水的来源和加工处理的次数，服务的顾客类别，人口密度，工资率，管制是松弛还是严格的，以及公营公用事业公司是否由管理私营企业的同一机构管理。他的回归分析结果显示，公有权的净效应是平均减少可变成本24%。他的数据还显示，所有权形式并不影响成本曲线的斜率。换言之，私营和公营企业的平均可变成本曲线和边际成本曲线的斜率一致，但公营企业的成本曲线却比私营企业的成本曲线位置低。虽然布鲁金克没有提供公营企业面临实现X效率的更大压力的证据，但他的数据与他的假设是一致的。

7.3.3 泰勒

威廉·泰勒(William Tyler)用类似于前述蒂默的方法对1971年巴西的16家塑料企业和22家钢铁企业的行业生产边界进行了估计。当时这两个行业都享受着抵御外国竞争的大量有效保护。因此人们会预料存在X低效率的企业可以维持经营。泰勒的塑料行业样本中包括11家国内所有和5家外国所有的企业，钢铁行业中包括15家国内所有(其中5家为公有)和7家外国所有的企业。

在两种行业中都存在对行业生产边界的较大偏离。塑料行业中16家企业的平均效率等级为0.48，而5家外国企业的等级为0.58。因此，如果每个塑料企业都具有本行业效率最高企业那样的(X)效率，那么在投入同样要素的条件下他们的产出平均可以增加一倍。22家钢铁企业的平均效率等级为0.62，其中5家公有企业的平均数字为0.56，7家外国企业

的数字为0.72。1971年公有钢铁企业的生产量达到钢总产量的54%。泰勒对这些数据作出的解释之一如下:"这可以被视为存在大量X低效率的证据"。较之外国企业,国内和公有企业的X低效率更大。

7.3.4 吉利斯

马尔科姆·吉利斯(Malcolm Gillis,1982 年)在玻利维亚和印度尼西亚的公有采锡企业中也发现了大量存在的X低效率状况。然而,吉利斯的证据来自关于这些企业内部活动的第一手报告资料。他列举的X低效率例子有:企业连续6年拒绝采用已知的高效率技术,因为"……以前从来没有这么做过"。另一个例子发生在1973年减少低纯度锡的出口税之后。1973年在公有企业中低纯度锡占产出的45%,私有企业的这一比例为43%。一年以后私有企业低纯度锡的份额增加到94%,而公有企业则降到17%。吉利斯指出,原因在于公有企业无视新的关税鼓励措施,继续其以前的计划,相反私有企业却充分重视到这些鼓励政策。公有企业显然跌入了一个"惰性区域"。此外,公有企业还存在其他缺陷,如会计技术更多基于"事后的收入决定",而不是依靠为当前决策收集的信息;而且,缺乏得到授权的管理机构则会导致漫长的停工期。对基于绩效的报酬的严格限制,以及用模棱两可的词句所表述的"社会责任"也是导致成本超出最低水平和其他私有企业水平的原因。

吉利斯将这些例子与配置低效率的情况作了对照。导致后者发生的例子是,在印度尼西亚和玻利维亚实行能源价格

管制一般会导致能源密集型投资项目的出现，这是与未受补贴的世界能源价格作出比较后的结果。另一个例子是，人为高估本国货币的价值会促进有利于进口的消费和投资组合。在配置低效率的情况下，产出和投入组合沿生产边界线发生变化；而在X低效率的情形中，企业是在其边界以内进行生产。

7.3.5 安德森和弗朗茨

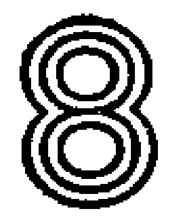
现在来谈谈汇率管制的问题，看看这种政策是如何影响X低效率的。1982年初标志着一个时期的结束，在该时期中墨西哥政府一直实行买入比索以人为地保持其在世界市场上高价的政策。1982年2月的汇率是27比索兑换1美元。2月19日，当墨西哥政府不再像以前那样实行支持其货币的政策时，比索对美元的比价开始下跌。在更多取决于货币市场供求关系的情况下，比价开始变成45比索兑换1美元。我和我的同事琼·安德森(Joan Anderson)开始着手调查这将怎样影响那些在墨西哥边界上实行计件工资制工人的劳动积极性或X低效率程度。这些工人对实际工资的变化特别敏感，因为他们对自己的工资收入规定了某种尺度。在诸如墨西卡利(位于圣迭戈东部约100英里，恰好靠近加利福尼亚州卡莱克西科的南部)这种"边境"城镇的居民中，1982年2月有约60%的消费支出是以美元计价的。因此，比索对美元的比价下跌约40%便意味着实际收入直接损失约25%。(45比索兑换1美元相当于1比索可以购买3.7美分。贬值后1比索只能买2.2美分，或者说比索按美元计其购买力下降了40%。)这个25%的

比率并不包括由此带来的额外通货膨胀，后者在年底已接近100%。

我们收集了墨西卡利两家纺织厂中每个实行计件工资制工人(共46名)的日产出率数据。这些数据的日期包括从1981年11月到1982年4月的22个星期。1981—1982年度的11—12月这一时期被作为“控制”时期。我们拟定对企业产品的需求、资本存量、以及制造特殊产品的难度为变量。计量结果发现，贬值的效果是每周平均增加产出15%。此外，其中一家企业在3月中旬增加了20%的工资，结果导致周产出又增加了8%。我们相信这两种结果是由不同原因造成的。对15%增长的最可能解释是，工人们试图减少源于贬值的实际收入下降。因此实际工资的下降使工人工作得更加努力。这意味着这些工人为了挣得收入是会尽心尽力的，就是说，他们试图保持某种目标收入水平。另外，由于工作小时是固定的，工人们不能随实际工资的下降减少工作小时数。因此，额外的产出并不出自更少的“闲暇”(即工作更多小时)。两家企业的总经理都认为这一解释准确地体现了他们员工对收入的传统流行态度。至于工资增加20%所导致的结果，我们认为那是因为贬值和增加工资造成了对管理者的新压力。这些企业拥有以美元计值的来自美国银行的贷款，但美国企业购买它们产品的合同却是以比索计值的。它们处于左右为难的境地。结果，按两家企业总经理的说法，就是为了增加每单位投入的产出，管理者要更加注重生产工艺的进步。在这种情况下，经济困难成为促使工人和管理者更加努力工作和提高X效率的催化剂。由于我们的数据仅延伸到贬值后的两个月左右，因此我们无法确定这种效应是否持久。

7.4 结论

本章涉及的研究似乎揭示了所有权形式中影响X效率的不同动力。虽然这些大多数的研究结果是X效率理论可以预料的，但其中的一些，例如布鲁金克的研究却(隐含地)提出了控制促使雇员追求X效率行为的所有压力的重要问题。这一问题在第9章会再讨论。下一章里我们将考察关于市场支配力及其对X效率影响的研究。



经验证据：市场结构

8.1 导言

我们关于X效率理论的经验证据方面的最后一章将涉及市场结构，或者说是厂商或行业面临的产品市场竞争程度的问题。本章也将考察三组文章。第一组重点讨论产出和投入比率，并以下列三种方式中的某一种衡量X低效率：(1)以相对产出(实际产出比可能的产出)来衡量；(2)以偏离其成本极小化价值的实际投入比率来衡量；以及(3)以期望的、实际的和利润极大化的生产能力利用率之间的比较来衡量。第二组重点讨论利润，而第三组则分析固定价格行为的几种X效率效应。

8.2 产出和投入比率

对许多高失业率和资本密集型技术的采用共存的欠发达国家来说，选择适用技术显得尤其重要。适用技术以及价格配置低效率和X(低)效率间的区别在图8-1中作了简略回

顾。在等成本曲线为CC′，等产量曲线为Q_1的情况下，企业如果在A点生产就能达到价格配置的高效率，因为在该点相对要素价格(W/R)等于要素的相对边际产品(MP_1/MP_2)。在A点厂商也实现了X效率，因为它的产出Q_1是在等产量线或生产边界上，即它在生产Q_1时使用了极小数量的投入要素。

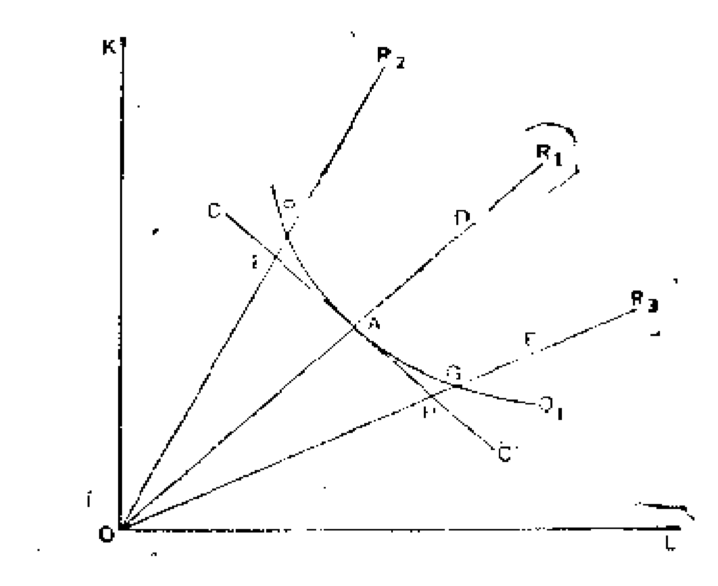


图 8-1　适用技术，价格配置低效率和X低效率

然而，在B点厂商则处于价格配置低效率状态，因为它虽处于生产边界，但却没有达到相对要素价格与相对要素边际生产率的相等。我们可以视OB/OF为价格配置低效率的尺度。另一方面，在D点上厂商达到了价格配置效率但却是X低效率的，因为虽然它使用了正确的要素组合(由于A和D点位于同一"射线"上，所以它们表示同样的资本/劳动比率)，但它却使用了比产出Q_1技术上所需的更多投入。D点上的X低效率程度是OD/OA。在E点上厂商既是价格配置低效率的(OG/OH)，又是X低效率的(OE/OG)。

8.2.1 勒克劳

价格(配置)低效率和X低效率现象在欠发达国家中有多普遍?唐·勒克劳(Don Lecraw,1977年,1979年)以1962—1974年间泰国12个四位数产值行业中的400家灯具制造厂为样本,对这种类型的低效率进行了估计。样本中的所有企业在这段时期内都刚刚在泰国进行了大幅度的新投资。因此,一些数据是基于计划的——事前——价值,即资本、劳动和投资中的计划利润。用到的这些和其他数据是通过对每家企业的采访获得的,其中对200家企业使用了一份全面的调查表,而对余下的200家企业则使用了一份较简短的调查表。勒克劳对X低效率的测算是通过估计一个固定替代弹性(CES)的生产函数获得的。该函数使用了一个非线性的形式,它不要求采取利润极大化的行为。一旦得到该CES生产函数的参数,就能对下列各项进行比较:实际的和可能的产出,实际的和成本极小化的资本/劳动(KL)比率(它要求掌握投入要素的价格),以及实际的和成本极小化的总成本水平。

勒克劳对CES生产函数运用与蒂默和泰勒相似的方法(他们采用的是科布—道格拉斯函数形式)估计了每家厂商的可能产出(Q^*),并将它与实际产出(Q)相比较。他把这一比率Q/Q^*叫作D_1。接着他把行业中所有厂商的这一数据累计起来,得出全行业的D_1值。小于1.0的D_1值表示厂商在一定的投入下没有获得最大可能的产出。在图8-1中这反映出厂商位于诸如D这样的点。因为D点表示厂商使用的投入超过了为获得一定数量的产出在技术上所需的数量,所以它也必

然表示厂商没能从D点代表的投入中获得最大可能的产出。因此小于 1.0 的D_1值可以衡量技术(X)低效率的程度。

价格配置低效率可以通过比较实际的资本/劳动比率(K/L)和使成本极小化的资本/劳动比率(K^*/L^*)来估计。他把这一比率(K/L)/(K^*/L^*)叫作D_2。如果D_2大于1.0，那么厂商的资本/劳动比率就是过分资本密集化了，即它的实际K/L比率超过了(在给定投入要素价格的情况下)使成本极小化的比率。回想一下，由于这些数据是事前决定的，所以D_2值不等于 1.0 并不表示投入要素价格变化引起的价格配置低效率。正因为这些数据是事前决定的，任何偏离成本极小化的投入要素比率的情况一定是由其他原因引起的。在图 8-1中，产生的价格配置低效率是由OB/OF这个比率衡量的。如果D_2的值小于 1.0，那么K/L 比率就是过分劳动密集化了。在图 8-1 中，过分的劳动密集度导致的价格配置低效率是由OG/OH这个比率表示的。

最后，价格配置低效率和X低效率的综合情况可以用所使用技术带来的实际总成本与如果选择最有效技术所招致的最小总成本之比来衡量。他把这一比率叫作D_3。如果一家企业选择适用技术(图 8-1 中的A点)，那么D_1、D_2和D_3的值都为 1.0。

事实是，泰国所有行业中资本密集度最高的企业较之资本密集度最低的企业，其K/L 比率是后者的 4 倍，这说明泰国可以对使用何种K/L 比率进行选择。D_1、D_2和D_3的值也表明价格配置低效率和X低效率都大量存在。例如，12种行业的D_1平均值是0.74，即从平均来看，每家企业的产出都比在现有投入下它应获得的产出少约25%。所有企业中的最低值是

0.44。行业平均值的范围在0.62到0.88之间。

12种行业的D_2平均值是1.71，表明这些行业总体上使用的是资本密集型技术。所有企业中最低值为0.38，最高为3.75。行业平均值范围是在1.25到2.02之间。12种行业的D_3平均值是1.60，即成本比技术上能达到的最小水平高出60%。所有企业中的最大值是2.42，而行业平均值范围在1.41到1.82之间。

然后，勒克劳试图用一系列(自)变量来解释D_1、D_2和D_3的值。这些变量包括：进行投资时预计的投资利润；反映产品市场竞争程度的行业中企业数目；企业是所有者管理还是由所雇佣的经理管理；企业当地所有的程度；以及经理在欠发达国家中是否有经验。

他对D_1的回归分析结果显示，下列因素中的每一个都会增加D_1的值，即让企业接近其生产边界或增加X效率：较低的预期利润，较多数量的竞争企业，企业的经理并非业主，以及经理具有在欠发达国家中的经验。他对D_2的分析结果表明下列因素会增加D_2的值，即导致使用资本密集度过高的技术：较高的预期利润，较少的产品市场竞争，企业是所有者管理的，以及企业经理缺乏在欠发达国家中的经验。毫不奇怪，对D_3的分析结果与D_2几乎是一致的，它表明下列因素会增加D_3的值，即增加超出最小水平的总成本：较高的预期利润，较少的产品市场竞争，企业由所有者管理，以及企业的经理缺乏在欠发达国家中的经验。分析结果还显示，当地所有权会减少D_1的值而增加D_2、D_3的值，但这一结果并不具有统计显著性。

企业数量增加的效应，即更具竞争性的环境会增加X效

率这一现象是不奇怪的。较低的预期利润会增加X效率的结果也是如此：较低的预期利润会迫使企业雇员走出他们的惰性区域，要求他们更关注成本极小化和使产出接近生产边界。出乎X效率理论预测之外的结果是，所有者管理的企业X效率较低。勒克劳对这一结果的解释来自于他对公司经理们的采访。所有者兼经理看来是对利润感兴趣的，因而是注重X效率的；但他们同样对从使用现代西方技术获得满足感兴趣，而这会导致一些成本方面的不利。这一现象被称作"工程师式的满足"，该术语是由路易斯·韦尔斯（Louis Wells，1973年）创造的。因此，虽然所有者感兴趣的不仅仅是利润，但被雇佣的经理并不知道这一点，他们倾向于假定所有者的兴趣仅限于追求利润极大化。事实上，按照勒克劳的看法，"缺席的"所有者对工程师式满足的兴趣要小于所有者兼经理。尽管被雇佣的经理会获得工程师式的满足，他的经营绩效却是由对其他东西感兴趣的缺席所有者来评判的。总的来说，该研究资料说明从提高X效率中获得的利益是巨大的。

[关于在欠发达和较发达国家中投入替代、技术变革和X低效率之间关系的理论框架和经验证据是由申(1973年、1984年)提供的。]

再论生产能力利用率

唐·勒克劳不仅估计了竞争、所有权形式、预期利润和其他变量对(Q//Q*)，以及对K/L比率和总成本对其极小值偏离的影响，而且也估计了它们对生产能力利用率的影响。勒克劳对生产能力利用率的衡量尺度是工厂在每年8760小时中运转的百分比。（8760这一数字代表一年的小时总数。）这

项研究的目的是为了理解上述讨论的要素是如何影响期望的生产能力利用率(D)，从而可以将它与实际的生产能力利用率(A)和利润极大化的生产能力利用率(P)作比较。在整个200家企业的样本中，期望的、实际的和利润极大化的生产能力利用率分别是30%、27%和67%。

然后，勒克劳将(P-D)这一差异作为衡量X低效率的尺度。他指出，诸如技术、投入价格、第二和第三班的工资差别以及产出率等"经济"变量非常适用于解释利润极大化的生产能力利用率；但它们对解释期望的生产能力利用率却并不那么有效。为了解释D和(P-D)，勒克劳使用了X效率理论和许多上述讨论的变量。他的回归分析结果显示，在"低"利润企业(那些股本收益少于25%的企业)中，当一班制经营的预期利润较高，而且企业是由泰国人或发达国家的公民所有时，D就较低。在这些企业中，当一班制经营的预期利润较高时，(P-D)较大。在"高"利润企业(股本收益大于25%)中，当预期利润较高，当该行业中企业数量较少——即竞争较少，当多班制经营的风险较大，当企业由所有者管理，以及当企业为泰国人或发达国家公民所有时，D就较低。在这些企业中，当预期利润较高，当该行业中有较多企业，当多班制经营的风险较大，当企业是所有者管理，以及当企业由泰国人或发达国家公民所有时，(P-D)也就较大。

另外，"非经济"变量在高利润企业中对D和(P-D)所起的作用比在低利润企业中更大。因为赢利高，这些企业似乎就能使这些因素在其决策中更具影响力。从我们对它其他方面的研究来看，下述结论并非出人意料，即当预期利润较高而竞争较少时，对利润极大化的偏离也较高。

8.2.2 卡尔松

布·卡尔松（Bo Carlsson, 1972年）也对竞争对X低效率（Q/Q*）的影响进行了估计，他使用了类似于蒂默、泰勒和勒克劳所用的测量方法，其样本包括1968年瑞典的26种制造行业中的企业数据。他对每个行业都估计了其科布—道格拉斯边界生产函数，其中产出以所增价值来计量。资本是交替以下列各项中的任一种来估量的：生产中使用的马力容量，消耗的电力，厂房、机器和设备的火灾保险值，以及除厂房外的火灾保险值。劳动力是以工人人数衡量的，它又分为技术人员、领班和一线工人。

卡尔松使用上述变量估计了边界生产函数，并从中导出每个行业的（Q/Q*）这一比率。然后，他把每个行业的（Q/Q*）作为因变量用于普通最小二乘法的回归分析中，以便解释行业间的差异。他选择的自变量是用来检验其主要假说的，即缺乏竞争压力是产生X低效率的主要原因。因此，他的自变量包括下列各项：行业中的4家企业集中率（4FCR）①，名义关税率，出口/产出比率，以及进口/消费比率。因为规模经济预期会提高效率，所以为了区分X低效率和规模经济两种效应，卡尔松使用了两种4FCR尺度。他的两种尺度分别是，未加权的全行业比率（CRu）和行业内主要商品类别的加权比率（CRw）。余下的三个变量则反映外国的竞争。

他的自变量中也包括行业中所生产的商品数目，这是衡

① 原文是four firm concentration ratio，指某一产业部门中最大4家企业在该部门总产值中所占的份额，它表示该部门生产的集中程度。——译者

量产出异质性的尺度。自变量中也包括每个行业中的观测单位数目(从 12 到154)，这一调整是为了对应只有这么多企业能在其生产边界上运转这一事实。因此，在偏离生产边界的分布情况已定的条件下，具有较多观测单位的行业其X效率会显得较低。自变量中还包括从估计边界生产函数得到的系数总和，以便为规模经济对 4FCR 尺度影响提供另一种修正。

最后，他还包括了一个重要变量，即乘以资本/产出比率的1968年行业资产的变化，它被用作资本利用率的代理变量。然而，它也可用作新投资效率的代理变量，即一些企业和(或)行业显得低效率是否仅因为它们仍处于"适应"新投资的状态？在前一章格雷戈里和詹姆斯的论文中这一问题已经提到过。

然后，卡尔松进行了几项回归分析，以便包括他的每一个4FCR 变量和几个衡量资本的尺度。他的分析结果可以概括如下：首先，两种 4FCR 尺度都是正的，且具有统计显著性。他对此的解释是，在像瑞典这么个较小和开放的经济体中，4FCR 更多地是反映规模经济而不是市场支配力，因此提高行业集中率看来会提高行业效率。第二，较高的名义关税率会降低行业的效率比率。第三，由进口/消费比率或出口/产出比率衡量的外贸参与水平对行业效率的影响并不大。最后，衡量新投资对效率影响的变量为负，但在统计上并不显著。

他试图探究国外竞争对效率影响的另一种方法是用相应"技术"要素调整每个行业的效率指数。这些要素是商品数目、观测单位数目和资本利用率。完成这些工作后他指出，26种行业的平均行业效率指数是0.78，变动范围在 0.9053 (机械制造)和 0,5770 (印刷工厂)之间。另外，效率最低的10种行

业通过较高的关税和其他对进口方面(纺织品和服装、酿酒、乳品业)的限制,或由于语言障碍(印刷工厂),得到了免受国外竞争的大力保护。另一方面,3个效率最高的行业——机械制造、汽车制造和船舶建造——位于瑞典制造业产品的最大出口者之列:它们的出口份额各占54%、43%和69%。这些结果与卡尔松的主要假说是一致的:"缺乏竞争是产生X低效率的理论上的主要原因。"

8.2.3 怀特

劳伦斯·怀特(Lawrence White,1979年)研究了1967—1968年间巴基斯坦的31种行业中国内和国外竞争、以及工程师式满足对资本/劳动比率的影响。怀特对工程师式满足效应的考察是通过假定巴基斯坦企业家希望仿效美国企业家的资本/劳动(技术)选择来进行的。为了检验这一点,尤其是美国资本/劳动比率的变化是否影响巴基斯坦的这一比率,怀特使用了在巴基斯坦和美国具有相似定义的31种行业的数据。为了检验竞争对资本/劳动比率的影响,他使用了巴基斯坦的4家企业集中率和巴基斯坦行业产出中的出口百分比两个变量。怀特又把后一变量一分为二,即至少出口其产出10%的行业被归入面临国外竞争压力一类,而出口少于10%的行业则被认为没有受到国外竞争。(10%当然是个有些任意确定的标准。)他的另一个自变量是个虚拟变量,以辨别该行业是化学工业还是加工型行业。两种行业对投入的要求都比较严格,因此其投入要素互相替代的可能性较小。资本/劳动比率是以三种不同方式定义的,每一种方式使用了对劳动的不同

定义：(1)总劳动，(2)生产劳动，以及(3)非生产劳动。

他使用的统计方法是普通最小二乘法。以三种不同的资本/劳动比率形式进行回归分析的结果如下：首先，美国的较高资本/劳动比率导致了巴基斯坦的较高资本/劳动比率。然而，随着美国该比率的上升，巴基斯坦的比率相对显得越来越低。因此他推断，工程师式的满足的确存在于巴基斯坦企业家的样本中。第二，当4家企业集中率上升时，巴基斯坦的资本/劳动比率也会上升。第三，当行业产出中的出口百分比不到10%时，巴基斯坦的资本/劳动比率较高。这些结果与他的假说是一致的，即竞争迫使企业更具X效率，或更注意成本。在巴基斯坦的例子中，注意成本意味着使用更多劳动密集型的技术。虽然怀特的数据与勒克劳和卡尔松所使用的相一致，但读者必须意识到的是，与勒克劳的数据不同，怀特的数据不能告诉我们较高的资本/劳动比率是否反映了价格配置和(或)X低效率。按图8-1来看，巴基斯坦的企业可能位于G点，但正在向A点移动，因为它们希望仿效美国人。或者，巴基斯坦企业可能在A点，但正向B点移动。其他可能性当然也存在。

8.2.4 伯格斯曼

另一项关于市场支配力、针对国外竞争的保护和X低效率之间关系的研究是由乔尔·伯格斯曼(Joel Pergsman，1974年)进行的 在这个例子中，伯格斯曼对巴西、马来西亚、墨西哥、挪威、巴基斯坦和菲律宾的占国民生产总值(GNP)一定百分比的市场配置低效率和X低效率(加上"垄断收益")进行

了估计。在所有情况下保护都招致成本，因为它使国产商品的生产成本高于不受保护情况下的成本。然而，当国内企业可能的最小成本超过国外企业时，保护就产生市场配置低效率。伯格斯曼将这些称作“不可避免的”高成本。通过保护国内企业不受国外竞争，外国产品的价格被提高了，消费者则选择了与自由贸易条件下不同的消费组合。

其次，当国内生产成本可以降低，但由于控制成本的压力被那些贸易壁垒减轻而没有降低时，保护会导致X低效率。他称这些高成本是“可以避免的”。最后，当尽管存在保护但国内和国外成本仍十分接近，而保护也仍被用来提高工资和(或)利润时，保护就会导致“垄断收益”。伯格斯曼承认要区分X低效率和垄断收益是十分困难的。因此在估量分析中，他把两者归并在一起。

他的估计结果是，上述6个国家属于市场配置低效率的保护成本(作为国民生产总值的一个百分比)各为0.3%、-1.2%、0.3%、-0.2%、0.5%和1.0%。另一方面，X低效率的成本(加上垄断收益)各为6.8%、0.4%、2.2%、2.0%、5.4%和2.6%。与莱宾斯坦原先对X低效率的论述一样，伯格斯曼发现X低效率比市场配置低效率大得多。

然后，伯格斯曼比较了三种类型经济体的结果：(1)以马来西亚、挪威为代表的小型和比较开放的经济体；(2)以巴西、墨西哥为代表的大型和相对工业化的经济体；以及(3)以巴基斯坦和菲律宾为代表的小型的、受到大量保护的经济体。这三类群体中市场配置低效率的平均水平各为-0.7%、0.3%和0.75%，X低效率的平均水平各为1.2%、4.5%和4.0%。小型和开放的经济体显示出较小效应的结果是并不令人奇怪

的。X低效率大于配置低效率的状况与本书其他部分公布的结果是一致的。但正如伯格斯曼所说的，从假定企业是追求成本极小化者模型的逻辑来看，保护导致X低效率效应这个结果是令人奇怪的。他说，"我发现，每一种这类估计都假定，可归因于保护的低效率仅仅是保护相对价格的效应所引致的配置不当的结果。这一假定通常是明确的，而有时又是隐含的，但却总是存在的：即保护导致了福利损失，因为它致使一个国家去生产与自由贸易条件下不同的产品组合，而且它减少了消费者剩余。"显然，他的数据与这种假定不一致。

8.2.5 佩奇

佩奇(*Page*, 1980年)估计了1972—1973年期间加纳的伐木、锯木和家具制造业中的X低效率。佩奇使用了类似于蒂默和卡尔松的计量技术，结果是三个行业的平均效率等级各为0.709、0.710和0.743。如果使用一种更为保守的技术，把它们当作一个回归方程中的因变量，则这些数字会是0.920。他的因变量包括：每个主管人监管的生产工人数目，经理的培训水平，企业年龄，经理中外籍人员的百分比，以及企业是由加纳人还是其他人所有。他的结果显示，管理培训质量对提高X效率具有重大的影响，拥有较大比例的外籍经理也是如此。但外国所有并不提高X效率。

8.2.6 安德森和弗朗茨

我和同事琼·安德森(1985年)估计了1982年墨西哥蒂华

纳和墨西卡利两地的24家纺织企业的行业生产边界。整个样本的平均效率等级是 0.469 。以这些等级为因变量的研究结果显示，墨西哥人拥有所有权会减少X效率。

8.3 利润和X效率

利用(价格—成本)差额和其他的利润衡量标准(如利润占销售额的百分比，或净收入占销售或资产的百分比)也可以对市场支配力和X低效率的关系进行估量。

8.3.1 韦斯和帕斯科

韦斯和帕斯科(Weiss and Pascce,1985年)认为，也许检验X效率理论的最好方法之一，是在某个给定行业中估计行业集中程度(自变量)和各企业的生产率方差之间的关系。另一种类似的检验是以利润代替生产率。两位作者利用联邦贸易委员会的商业资料中关于1974—1976年间 233 种制造行业的数据，进行了这样一种检验。在其研究过程中，两人所作的检验对一个重要问题提供了一些证据，这个问题就是关于X低效率和规模经济(次优生产能力)各自对生产成本的独立影响。首先用图表来说明这个问题：X低效率和次优生产能力(以及过剩生产能力)在图 8-2 中被显示出来①。在长期平均总成本(ATC)曲线上，工厂1(短期平均成本 $SRAC_1$)是最优规模的厂商。工厂2(短期平均成本 $SRAC_2$) 是次优的厂商，

① 西格弗里德和惠勒(Siegfried and Wheeler, 1981年) 对区分X低效率、过剩生产能力和次优生产能力的问题提供了有用的论述。

因为它的规模不够大，无法获得本行业的所有规模经济成果。如果工厂2计划生产Q_2，那么它的次优生产能力由(ATC_2-ATC_1)来衡量。此外，如果厂商按$SRAC_{2'}$而不是按$SRAC_2$生产，那么与产量Q_2相联系的X低效率水平为(ATC_3-ATC_2)。另一方面，如果工厂2计划生产Q_1，那么它的次优生产能力是(ATC_4-ATC_1)。然而，Q_1对工厂2来说并不是个高效的产出率，因为一家更小的工厂可以以更低的平均成本生产Q_1。因此，使工厂2仅生产Q_1会招致比次优生产能力所带来的更高成本。这一附加成本被称为"过剩生产能力"。过剩生产能力，即没有使工厂按其长期平均成本的最低水平进行生产，它所提高的成本数量是(ATC_5-ATC_4)。最后，如果企业是按$SRAC_{2'}$而不是按$SRAC_2$生产的，那么与产量Q_1相联的X低效率就是(ATC_6-ATC_5)。

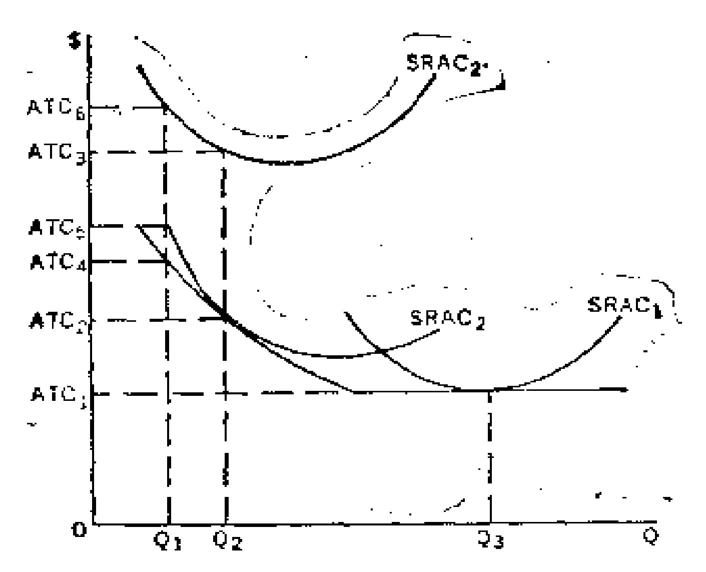


图8-2　X低效率、次优生产能力和过剩生产能力

让我们回到韦斯和帕斯科的分析中来，他们衡量利润的尺度是除以销售额或资产额的营业收入的方差和标准差。因

此，这种一个行业内利润率的离散可以作为X低效率的代理变量。原因在于，如果市场支配力使各企业避免了实现X效率的压力，那么可以预料比较高度集中的行业会表现出更多的行为模式，即会展现出更多不同程度的理性。

在第一组回归分析中使用的自变量是：经相对非竞争性的子市场调整后的4FCR；测量非竞争性子产品的市场细分程度(例如，凭医生处方出售的药品包括维生素、避孕丸和抗菌素)的赫芬达尔指数；产品运输的距离，这是市场细分的另一个代理变量；以及行业三年平均营业收入与销售额和资产额的比率。这些普通最小二乘法的回归分析的结果中包括与以前一致的结论，即较高的4FCR增加了行业间利润的离散度。可以假定，在更高度集中的行业内X低效率更大。

然而，这种结果也可能是下述事实导致的：即高度集中的行业中大企业的数量较少，而其所占的市场份额很大，同时大多数企业却只拥有很小的市场份额。在这种情况下，利润的方差之所以大可能是因为只有这些少数大企业赢利，而多数较小的企业获利要低得多。另一方面，集中程度较低的行业之所以显示出较小的利润离散程度，也可能是因为所有企业在规模和市场份额上更为平均。因此，以前公布的结果可能是由于企业规模的分布和市场份额，而不是由X低效率引起的。

要纠正这一点就必须区分X低效率的效应和源于次优生产能力的低效率的效应，这种低效率是因为企业太小，不能利用所有的规模经济效应而导致的。这里使用的方法是将其利润尺度按每个行业销售额的对数进行回归。这些回归中得到的残差接着便成了他们下一组——第二组——回归中的因变

量。在这组新回归中自变量与以前相同。这组回归的结果表明比较高度集中的行业中具有更大X低效率的证据要少得多。对此可作的一种解释是，利润的较大离散更多是由于次优生产能力的作用，即企业无法得利于规模经济的效应，以及由于市场份额的分配，而不是由于X低效率的作用。

接着，两人又指出，在233个行业中157个行业的销售量和利润呈正相关，而145个行业呈负相关。呈负相关的行业被认为处于非均衡状态。因此，在他们的第三组回归中，他们只将销售量和利润呈正相关的行业进行了与第二组相同的回归。结果再一次显示，较高的集中率会提高行业间利润的离散度。因此，在"正常"条件下，由于X低效率而不是由于次优生产能力的作用，较高的集中率会提高企业利润的离散程度。

然而，销售额和利润的正相关可以是规模经济的结果，或者说正是由于大企业能更有效地管理，它们才能扩大规模和获得更多的利润。也就是说，一些企业"较优"而另一些则"较差"。德姆塞茨(Demsetz, 1974年)和佩尔兹曼(Peltzman, 1977年)就持这种观点。为了检验这些相左的观点，两位作者使用了普通最小二乘法和一个联立体系(二段最小二乘法，其中利润和市场份额都既是因变量，又是自变量)，对利润水平(而不是其变差)和市场份额进行了回归，回归方程中包括行业的4FCR和其他几个自变量。他们的结果可以概括如下：首先，在普通最小二乘法回归中，较高的利润增加了市场份额，而较大的市场份额提高了利润。第二，较高的4FCR会减少利润，但统计上并不显著。

第三，在二段最小二乘法回归中，虽然较大的市场份额会提高利润，但较高的利润都会减少市场份额。较大的市场份

额会增加利润这一点是与规模经济的存在和利用相一致的。但是，较高的利润会降低市场份额这一情况却与德姆塞茨和佩尔兹曼的观点相矛盾，因为他们认为高效率的领头企业会扩大规模和获得更多利润。这个结果与一种更古典的理论相一致，这种理论预测利润会吸引新企业，从而减少现有企业的市场份额。第四，较高的4FCR会降低利润水平。这点与X效率理论一致，即同一种市场支配力既会促使企业提高其产品的价格，又会减轻其追求成本极小化的压力。因此，当成本增加而不是收益提高时，利润会降低。

韦斯和帕斯科总结说，第一和第三组回归是与在比较高度集中的行业内X低效率会增加这种观点相一致的。虽然这个结果部分是由规模经济因素导致的，但他们相信，在高度集中的行业内X低效率仍是一种活跃的力量。最后，他们对这点没有疑问，即按照定义，一些企业是较优的而另一些则较差。然而他们的研究显示，较差的企业最可能集中于比较高度集中的行业。

8.3.2 谢泼德

按威廉·谢泼德(William Shepherd, 1972年a)的说法，“在市场需求一定的条件下，新古典理论预计较大的市场份额……会使企业获得较高的利润率……这个函数可能是线性的；或者，在市场份额较大的同时X低效率也较大时……它就可能是曲线的”。谢泼德利用1956—1969年间(一个持续的经济增长时期)《幸福》杂志美国最大500家企业名录中200家企业的数据，对这些观点进行了检验。然后，他在第二个500

家最大企业名录中又挑出了一个含231家企业的样本。企业被排除出样本是由于下列原因中的任何一条，（1）它们的业务是高度多样化的(因此，一个为许多市场生产的企业避免了分析利润和市场份额关系的问题)；（2）它们在该期间是一个较大的合并企业中的一部分；（3）他们的销售额中提供给军方的部分占很高的百分比(越南战争时期)；（4）他们有一年以上未被列入《幸福》杂志500家最大企业的名录中；以及（5）企业或行业正经历着一个重大的不平衡时期。

整个样本被划分为若干子集。第一个样本子集包括210家企业，其中由于对数据有特别疑问而剔除了21家企业。另一个子集包含50家生产同质商品的老行业中的企业，目的是为了检验这样一种理论，即占统治地位企业的行为会日趋消极和保守，以至于其市场份额下降，利润的增长放慢或实际下降。还有一个子集包括181家在“年轻”行业中的企业。利润这一因变量是以税后净收入占股本(股本的帐面价值加上留存收益)的百分比来衡量的。除了企业的市场份额以外，其他自变量还包括4家企业集中率、进入市场的障碍程度、企业规模(以其净资产计)、销售额中广告支出的百分比、以及企业总收入的增长率。

谢泼德的数据显示，当企业具有较大市场份额时其利润也较高，但在一般情况下，这种增长的速度是逐步减慢的。换言之，两者的关系是倾向于曲线的，老行业中的企业尤其如此。其次他指出，企业规模和利润呈负相关，即大企业的赢利能力要比小企业差。关于这项研究结果他评论说，“这一负系数是不是X低效率导致的呢？看来没有别的解释比这更有说服力……企业规模可能增加市场支配力，但看来它也增加了

成本(至少在1956—1969年是如此)。"第三，谢泼德指出，在50家"老"行业集团的企业中，当企业的市场份额为27%时，利润达到极大值。相比之下，181家"年轻"行业中的企业是在其市场份额为75%时才达到极大利润的。谢泼德评论说，"区别在于，我们所看到的在市场份额超出27%以后赢利能力的下降，至少部分是由于相对X低效率的作用，这正是钢铁和肉类加工企业的证据所不断表明的。"

在第二项研究中，谢泼德(1972年b)使用了1963—1967年间 336 种行业的数据。他在这里使用的利润尺度是边际成本定价或价格-成本差额(即价格—边际成本/边际成本)的离差。他指出，虽然拥有较高的 4 家企业集中率的行业其价格-成本差额也较大，这一关系并不适用于33种生产同质产品的老行业子集。也就是说，这些行业市场集中度或支配力的提高并不带来利润的增加。第二，与市场集中度方面的研究结果相似，除了老行业中的企业以外，所有企业规模的扩大都带来利润的增加。谢泼德总结说，这些结果至少部分是由于伴随绝对规模和市场支配力增加而来的X低效率的增加。

8.3.3 食品系统研究小组

威斯康星大学食品系统研究小组(Food Systems Research Group, 1977年)的成员布鲁斯·马里恩(Bruce Marion)和他的几位同事，研究了1970—1974年间美国36个大城市中17家领头食品连锁商店的市场支配力与其价格和利润的关系。这17家食品连锁商店约占1974年美国连锁商店的52%和食品杂货店的6%。另外，17家食品连锁商店中的每一家，其中包括塞

夫韦商店、P&P超级市场、温—迪克思商店、幸福商店、食品交易商店、大联公司、客来公司和艾伯森公司等，都位居1974年全美最大的20家食品杂货联号商店之列。利润是以占销售额的百分比(PR/SA)来衡量的。

几位作者按不同水平的市场集中度和企业的市场份额(FMS)，比较了估计的"市场篮子"的价格指数和利润。4家企业的集中率(4FCR)使用了40%、50%、60%和70%这几种水平；企业的市场份额使用了10%、25%、40%和55%这几种水平。他们预料，在企业市场份额一定的情况下，市场集中度的提高会同时提高价格和利润。同样可以预料，当4家企业集中率一定时，企业市场份额的增加也会同时提高价格和利润。研究所得数据证实了这种一般关系。例如，给定FMS为10%，相对于40%、50%、60%和70%的各个4FCR水平的价格指数分别是100.0、101.0、103.0和105.3。同时，(PR/SA)分别是0.37%、0.99%、1.22%和1.28%。另一方面，在FMS为55%时，四种4FCR水平下的价格指数分别是103.6、104.5、106.5和108.9。同时，(PR/SA)分别是2.71%、3.33%、3.56%和3.62%。在给定一个4FCR的条件下，价格和利润随FMS单调增长的现象也可以在数据中得到证明。

显然，当企业市场份额和行业集中率都上升时，价格和利润也都上升。然而，数据也显示，价格比利润上升得快。在FMS为10%和4FCR为40%时，价格指数是100.0，(PR/SA)是0.37%。在FMS为55%和4FCR为70%时，价格指数是108.9，而(PR/SA)是3.62%。换句话说，价格上升了8.9%，而利润仅上升了3.25%。如果在这两个极端的例子中，每一美元销售额的营业成本相同的话，价格和利润应当上升相同

的百分比。然而数据显示，增加的利润占价格上升部分的36.5%(3.25/8.90)。余下的63.5%被每一美元销售额的营业费用的上升所承担。换言之，市场支配力的提高所增加的成本大于它所增加的价格，因此利润增长的速度是逐渐放慢的。几位作者相信，X效率理论至少为这些数据提供了一个很好的解释：尤其是，市场支配力的提高使企业可以提高价格，结果这会鼓励企业放任其成本出现不必要的上升。价格和利润上升之间的差别是X低效率的一种量度。

8.3.4 麦克费特里奇

唐纳德·麦克费特里奇(Donald McFetridge,1973年)是以下面这段话开始其对1965—1969年间43种加拿大行业的研究的："这里运用的方法只对市场的运行状况提供了部分分析，因为注意力仅仅集中于静态的配置效率，而诸如成本极小化等市场运行的其他方面……被忽略了。与配置效率相比，这些方面很可能会对社会福利产生更为深远的影响。遗憾的是，它们不太容易进行分析，因此留给了未来的研究者。"麦克费特里奇衡量(P—C)的尺度是毛利率，即(P—AVC) Q/(VA)；其中，P是最终产品的价格，AVC是平均可变成本，Q是产量，而VA是增加值。假定成本不变，那么这个式子就是衡量作为增加值的一个百分比的价格和边际成本离差。然后他对(P—C)进行了回归，回归中包括下列各种变量：企业和工厂的集中率尺度，资产规模，增长率，该行业是生产消费品还是生产资料，衡量产品差别的广告强度，以及实际关税率。麦克费特里奇预料，较高的集中率和实际关税会增加(P—C)。后

者由于减少了国外竞争，使国内企业在提价时较少担心失去其竞争地位。他的分析结果显示，较高的关税率导致了(P—C)，但并不具备统计显著性。麦克费特里奇对这一令人惊讶的结果是这样评论的："可能的原因是，实际关税率与次优生产能力和超额成本有关，而不是与超额利润有关。"

麦克费特里奇是从研究市场支配力对配置(低)效率的影响开始其分析的，但他在总结其研究时却说道："要评价各种市场结构就需要了解配置低效率与某一给定的市场结构的相关程度。一个完整的评价也需要考虑……与各个市场结构有关的……X 低效率的程度。"

8.3.5 卡特拉克

霍米·卡特拉克(Homi Katrak，1980年)在他对印度制造业的研究中也使用了价格-成本差额。卡特拉克衡量(P—C)的尺度是[(P—W—M)/P]，其中P是产品价格，W是工资，M是原料成本。由于规模报酬不变，它相当于[(P—MC)/P]。卡特拉克试图不仅使用卖方集中率和资本密集度，而且使用衡量对外贸易的变量，来解释行业间(P—C)的变差。集中程度以该行业中 4 家最大工厂占行业资产的百分比来度量。资本密集度则以资本/劳动比率和资本/产出比率两种尺度来衡量。对外贸易变量包括占行业产出一定百分比的进口、占行业产出一定百分比的出口、以及对进口商品和投入物征收的"隐含"关税率。(隐含关税率不仅包括关税率本身，也包括其他对进口的非关税限制。)卡特拉克的理论预测，较高水平的卖方集中率，占行业产出较小百分比的出口和进口会提高

(P—C)，因为它们都增加了国内企业串谋和提价的能力。

卡特拉克的回归分析结果显示，(P—C)在较高的隐含关税的保护下会增加。他的结果还显示，(P—C)的增加要到集中率达到50%时才停止，随后便会下降。他问，后一种结果是否是集中率超过50%以后价格降低和成本升高所引起的。如果X低效率在这些较高的集中率水平下变得十分普遍时，那么按照卡特拉克的说法，这一点便可以从较高的单位劳动(w)和单位原料(m)成本，尤其是从前者中得到证明。而且，如果(w/m)充分增加的话，那便会抵消任何降价，因而便减少了(P—C)。

因此，卡特拉克按卖方集中率和资本密集度(资本/劳动)对(w/m)进行了回归。回归结果显示，(w/m)随卖方集中率的上升而上升。具体而言，卖方集中率上升1%，则(w/m)上升0.5%。因为卖方集中率的范围在0.1到0.9之间，卡特拉克断定，当集中率超过50%时，(w/m)的潜在增加可以大到足以导致(P—C)的下降。换言之，(P—C)的减少可能是由于X低效率成本的增加，而不是因为价格的降低。[卡特拉克又指出，(w/m)的增加在集中率更高的行业内也可能是工会诱致的工资提高的结果。]

8.4 固定价格行为

8.4.1 生产率差别和卡特尔

弗雷德里克·谢雷尔(Frederick Scherer)和其同事在他们关于工厂规模和多厂经营经济学的研究中，讨论了针对国

际竞争的保护对X效率的影响。这项研究的一个方面，就是比较美国、英国、加拿大、瑞典、法国和西德的8种行业中企业的劳动生产率。几位作者报告说，在这8种行业中，如果把美国平均劳动生产率的值定为1.0的话，那么其他五国劳动生产率的平均值为0.47，生产率差别中包含了相对的X效率。作者们指出，"我们的采访提供了大量的定性证据，证明纯粹的X低效率是造成生产率差别的重要原因……英国一些行业中的管理人员承认，生产率徘徊在低水平上，因为卡特尔协议造成了安于现状的态度。在卡特尔解体并随之出现竞争以后，两例中的状况出现了巨大改观……一位经常旅行的美国油漆行业的官员报告说，在参观欧洲工厂时，他看见'许多人在毫无意义地闲荡'，这种情况尤其出现于一个在油漆制造者间具有卡特尔历史的国家"。①

美国的体育馆座位和结构钢行业

在美国的体育馆座位和结构钢行业内也有类似的协同定价情况。这些与谢雷尔的叙述和X效率理论一致的例子是W·布鲁斯·埃里克森(W. Bruce Erickson,1976年)提供的。然而，埃里克森没有在他的文章中明确提到X效率理论。在体育馆座位联合会领导下的体育馆座位定价行动开始于1944年，持续到1953年。1954年，在折叠式体育馆座位联合会的指导下又实行了新的固定价格协议。联合会由占该行业销售额90%的6家主要的体育馆座位制造商组成。该联合会从1954年4月起每月召开一次会议，直到1960年初经司法部的

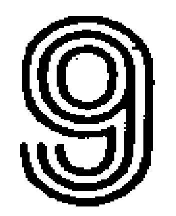
① 关于由据认为可能具有X效率效应的工会所导致的生产率提高可参见克拉克(Clark,1980年)的文章。

一次调查后它被解散为止。同多数卡特尔一样，除了需求下跌时期，这项协议一直是成功的。在串谋期以前和以后的价格平均分别为每英尺看台 6.43 美元和 6.95 美元。在串谋期间，平均价格为9.06美元，或者说比其他两个时期的价格高出30%到40%。串谋期的每英尺成本大约比串谋期前高60%。原因在于，价格串谋阻碍了创新和成本控制计划，并大大提高了工资，尤其是中高级管理人员的薪水。串谋期后的成本大约下跌了23%。

结构钢行业中的串谋定价是从 1950 年到 1962 年底实行的。这一时期的价格大约比此前和此后两年的价格高 25%。另外，这一时期的公司利润一般是美国同期所有行业平均值的两倍。最后，由于这一时期技术变革和生产率增长的迟缓，生产结构钢的成本提高了至少10%。当串谋期后两家新加入的竞争者以牺牲老牌的串谋企业为代价扩大其市场份额时，这些串谋者的惰性就明显地反映出来。

8.5 结论

这一章对许多国家的多种问题提供了大量研究。别人提供的证据，以及在一些例子中我们自己对其研究结果的解释，都是与X效率理论相一致的。我们现在转入对这些证据的另一种解释，这是由X效率理论的批评者给出的。



X效率：它的批评者和答复

9.1 导言

有关X效率的著作文献，不论是它的理论方面还是经验方面，自莱宾斯坦1966年的论文发表以来，都已是浩如烟海了。事实上，在1969—1980年之间，他于1966年发表的论文是《社会科学引用索引》①中第3篇最常被参考的论文。而且，由于有关X效率的著作已经日益发展和成熟，这一理论已引来可资识别的诸多不同"学派"的批评。在本章中，我将探究这些批评并提出一个回答。这一研究将得助于图9-1的运用。

图9-1中，竞争性厂商生产产量为Q_c，要价P_c，拥有长期成本C_c。垄断者将生产产量Q_m，要价P_m，成本为C_m。因垄断力量造成的市场配置损失则为面积ABC。面积P_mABP_c有4个组成部分：面积1表示利润，面积2表示较高的工资和薪水，面积3表示工作（"额外补贴"），但也将包括因诸如次优

① 《社会科学引用索引》包括一个被称为"近期期刊目次"的刊物，人们可从中复阅常被参考的文章。他们复阅的莱宾斯坦有关X效率理论的文章刊登在1982年8月16日第33期的第20页上。

生产能力和(或)超额生产能力等因素造成的较高成本。最后，面积 4 表示纯X低效率。然而，面积 2 和面积 3 可能也包含X低效率。如果因工资和薪水引起的较高成本不是"理性的"工资政策结果，这些较高成本就可能反映出X低效率。如果，例如额外补贴的成本超过这些补贴对其接受者的价值，面积 4 可能包含X低效率。(雇员可能宁愿选择现金，但税法则鼓励厂商给予补贴。)

我将运用图 9-1 详细阐述 4 种学派的批评，其中一些学派我们已经讨论过了。这些学派是：(1)寻租；(2)闲暇作为产出；(3)竞争下的管理效用；以及(4)产权。

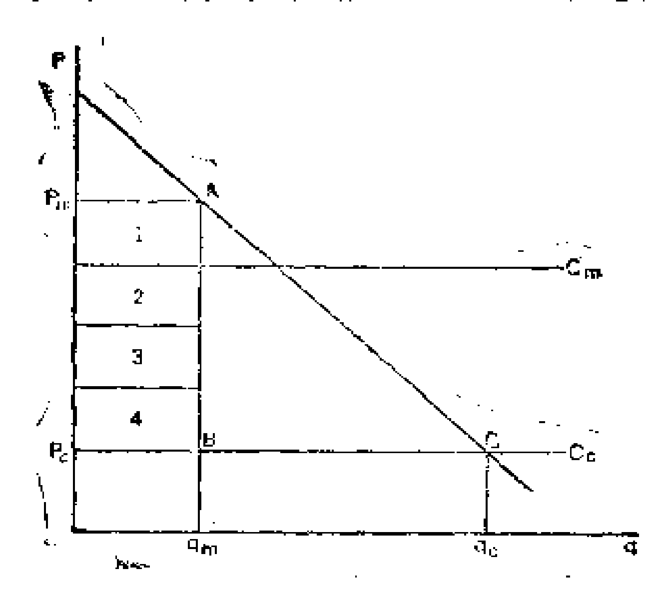


图9-1　划分垄断力量的较高成本：利润、工资、工作额外补贴和X低效率

9.2　寻租

我们已经在第 2 章讨论了寻租。概括地说：(1)塔洛克简单地假设，寻租支出体现于(观察到的)全部成本增加；

(2)克雷恩和扎德库希断言，X低效率是寻租支出的一个"残性"函数，这使寻租支出的增加(减少)(我将假设它可划归面积3)被X低效率等量大的减少(增加)所抵消。由于X效率带来的收益(损失)被寻租支出的增加(减少)所抵消，X(低)效率中的这些变化对净福利便无任何影响。(换言之，即使它存在，也是不重要的!)

9.2.1 沙普

戴维·沙普(David Schap,1985年)提出了一个论点，它类似于其他人士以寻租为题所作的论著。沙普的论点是以这样的假设为基础的：垄断的较高成本起因于寻租、租金消耗过程。他的论点在图9-2中得到说明。假定竞争性和垄断性的厂商成本是相同的，都为C_c，那么具有X效率的竞争性厂商和具有X效率的垄断性厂商的产出、价格和成本相应地分别为Q_c、P_c和C_c以及Q_m、P_m和C_c，损失的余额便是我们已经熟悉的面积P_mACP_c，其中面积ABC为市场配置低效率。沙普继而断言，面积P_mABP_c是因寻租和租金消耗过程而形成的，本身便排除了X低效率存在的可能性。

如果垄断者是X低效率的，它就将显示出在价格P_m'下的单位成本C_m和产出Q_m'。把X低效率的垄断性厂商与X低效率的竞争性厂商作一比较，就可说明损失的余额为面积$P_m'DCP_c$。其中，面积DFC为市场配置低效率，面积$GHFP_c$为垄断者生产Q_m'造成的较高成本，而面积$DHGP_m'$为因寻租和租金消耗过程造成的较高成本。

沙普然后运用图9-2显示了可归因于一个X低效率垄断

者的福利效应。首先，面积ABFJ表示因X低效率垄断者引起的附加市场配置低效率。我相信，沙普正确地指出，虽然莱宾斯坦利用生产成本论述了X低效率，作为低效率的一种非配置形式，这些较高的成本自身就产生了一个配置低效率。其次，面积$GHFP_c$表示具有X低效率的垄断者生产$Q_{m'}$的较高成本。再次，面积P_mJHG表示寻租支出。换言之，面积ABP_cGHJ是一种福利损失，不管垄断者是有X效率的还是X低效率的。如果垄断者是具有X效率的——ABP_cP_m，那么在寻租上支出的将是在X低效率垄断者条件下的一个市场配置损失（ABFJ）和一个X低效率（$GHFP_c$）。仅因X低效率垄断者造成的福利损失就是面积$P_{m'}DAP_m$。

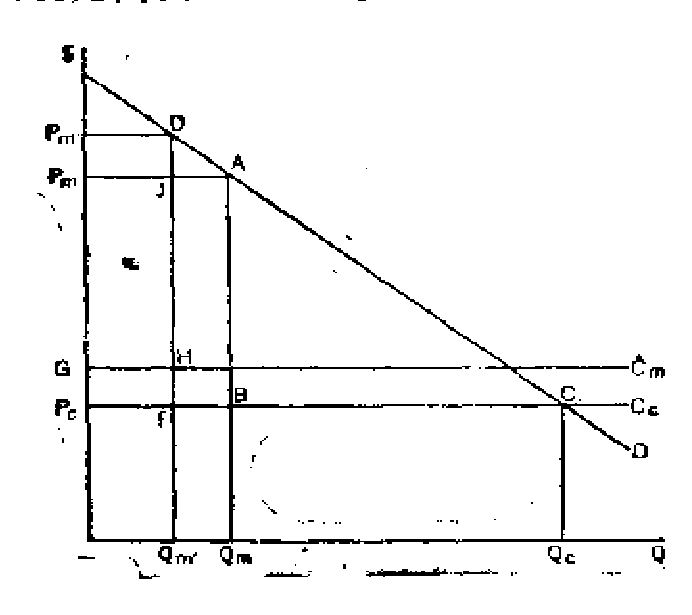


图 9-2 寻租、市场配置和X低效率

沙普还作了另一个重要的假设，即一个垄断性厂商的较高成本对该厂商来说是固有的，因而不可避免。在这一假设条件下，既然它不能被排除，面积$GHFP_c$——X低效率——就与生产者无关。结果，所有权权力对内在固有X低效率的垄

断者来说价值较小。理性的所有者因此较少从事寻租活动，它正好等同于X低效率所代表的较高成本。简而言之，我们转回到了克雷恩和扎德库希的观点。仅因X低效率造成的福利损失为面积$P_{m'}DJP_m$。

最后，如果X低效率的垄断者生产了Q_m而非$Q_{m'}$，那么面积ABC表示市场配置低效率，而面积ABP_cP_m表示寻租支出。由此得出X低效率的净福利成本为零。（换言之，如果你一开始就断言X低效率不存在，你就一直会得出结论：X低效率确实不存在！）

9.2.2 对寻租论者们的回答

塔洛克与新古典主义正统学派的分道扬镳在"精神"上与莱宾斯坦很相似。然而，这种相似的背离导致了不同表达的关注和理论。莱宾斯坦考察了作为成本和生产力变化根源的非极大化，这些变化在他关于X效率的最初论述中已作过研究和报道。塔洛克引用了部分这些相同的资料，但继续研究作为福利损失根源的寻租。莱宾斯坦看到成本处于最低水平之上，便称之为X低效率。塔洛克没有发现关于这些较高成本中的任何低效率，他猜测它们表示对政府造成的鼓励的一种理性反应。如此相同的出发点怎么会导致侧重点的如此截然不同，看来很有意思。同样十分有趣的是，比尔曼和托利森(Bierman and Tollison，1970年)引用了莱宾斯坦的原始论文，支持塔洛克关于寻租的理论。

对寻租理论和X效率理论不易作比较。寻租是极大化行为的一种形式。寻租行为是有目的和方向的；它是对一个利

润刺激极有策略的反应。进而言之，厂商被认为是一个生产函数。也就是说，所有雇员都是有方向的，总该通过管理去实现一个可以实现的共同目标。因此，厂商在寻租活动上消耗稀缺资源，直至其边际利润等于边际成本；这样，给定厂商的寻租支出，该厂商便是一个成本极小者。

最后一点是，厂商是否真的严格地遵循了寻租理论所指出的去计算其寻租预算。（这个问题已在第5章的附录中论及。）相信厂商的预算并非如该理论所指出的那样被仔细进行计算的一个理由是，厂商"寻租预算"的若干部分，也许是一个重要部分，被表示成是因同业公会引起的。换言之，厂商常常不对寻租支出的利润和成本进行边际计算。在这些情况下，厂商只是简单地付钱加入一个联盟，让它去完成这些活动。这些活动常常具有一个很普通的特征，比"手臂转动"更加信息化。①

9.3 闲暇作为产出

已有好几个作者运用这一方法，最著名的是乔治·施蒂格勒（George Stigler）。关于图9-1，这一方法或否定纯X低效率（面积4）存在，或断言面积3表示在职闲暇，因而完全是一种转移。后者的论点说明在职闲暇至少对其全部成本来说是有价值的。

这种方法以一个厂商生产两种类型商品的观念为基础。首先，厂商生产要出售的商品，即生产人们普遍认为是厂商

① 我是在同太平洋煤气和电气公司副总裁约翰·丹尼尔森（John Danielson）的交谈中获悉这一知识的。

产出的商品。其次，厂商生产"不可交换的"商品，诸如在职闲暇、"平静的生活"、友谊、休息和娱乐以及其他在内容上相似的商品。这样，如果厂商是一个垄断者而可能造成的较高单位成本就代表雇员们的效用，它本身就是垄断的利润，而不是垄断的社会成本。假定理性的或极大化的行为则意味着个人将有助于"生产"这些不可交换的商品，直至边际利润正好等于边际成本。换言之，如果一个厂商因为雇员正接受在职闲暇而未将其成本极小化，那么，必定雇员们会选择在职闲暇，而不是他们故意放弃的利润(收入)。

对莱宾斯坦关于比极小成本更高的成本是低效率的一种形式的理论的一个反应是，这一理论只有在从在职闲暇中获得的效用被大打折扣时才能成立。但当然，如果人们选择闲暇，那它怎么会减少呢？因而，莱宾斯坦一定是根据自己提出的定义而非其他人的选择来随意地定义低效率的。

9.3.1 帕里什和恩格

帕里什和恩格(Parish and Ng,1972年)是首先提出这一论点总概貌的学者之一。他们陈述说，较高成本是配置低效率而不是X低效率的一种形式。帕里什和恩格说，如果企业家们理性地采取行动，那么X低效率就不会产生任何净社会成本。这是因为企业家们的闲暇是社会产品的部分，而不是社会成本的部分。社会成本只有在这个意义上产生，例如当闲暇的定价存在一些低效率时。一旦这种情况发生，那么以收入替代闲暇的人们就不承担其作决定的全部成本。按他们的观点来说，对闲暇免于征税是造成单位生产成本上升的价格或配置

低效率的最"明显"例子。

9.3.2 莱文、皮尔和小帕索尔

闲暇论点的运用在这以后还常常出现。如莱文(Levin,1974年)争辩道，商品和劳务的生产遵守物理学规律，尤其是，它严格遵守物质和能量"守恒"定理。莱文把这解释成，如果厂商生产的可交换商品比可能生产的少，那么，厂商一定生产较多的不可交换商品。总之，在一个固定的工作日里，能量的消耗必定有一定的形式，如果它不是以可交换商品的形式表示，就必定以不可交换商品的形式出现。当能量以不可交换商品的形式表示时，它采取的一个形式就是在职闲暇。

皮尔(Peel,1974年)提出这样的论点，如果生产水平使其自身每单位努力(惰性)水平的产出极大时，一支"懒惰"的员工队伍是具有X效率的。他进而相信，把实际产出视为极大化产出"在使用上更加方便"。另外，他认为当成本增加了雇员们的激励因素时，非成本极小化是与X效率相一致的。管理和红利方面的支出便是这种成本的两个例子。

小帕索尔(Pasour,Jr.,1982年)提出了相似的论点，他认为较少的产出和更多的闲暇并不必然是任何低效率形式的一个标志。而且，既然选择了闲暇，很明显它必定比人们放弃的收入更加值钱。最后，帕索尔将低效率行为定义为净收益机会尚未被注意到的结果。于是，低效率造成了市场的非均衡；因此，根据定义，所有低效率都是配置的。既然未被发现的机会是"全然主观的"，它们不可能由除个人以外的任何其他人发现。结论便是，当另一个人处于低效率时，没人可以确定。

9.3.3 马丁

马丁(Martin, 1978年)在国际贸易保护的效率效应中讨论了努力-闲暇的权衡问题。马丁从管理努力的方面给X效率下定义。较多保护的效应并不必然增加X低效率。效应是根据这种保护的收入和替代效应的相对强度而定。一个可能的说明便是，通过提高厂商的产品价格，保护的增加提高了闲暇的价格。努力程度的增加——替代效应——是人们可以预料的。另一方面，较高价格的收入效应将降低努力程度。马丁总结道，与其他因素相比，当收入效应比替代效应更有影响时，保护将增加X低效率。

马丁还告诫我们要注意到保护的静态效应和动态效应两个方面。这样，当保护可能在进口部门增加X低效率("暖日"效应)时，它也可能在出口部门减少X低效率("寒雨"效应)。马丁认为伯格斯曼的研究(我们已在第7章中讨论过)是静态的，因而可能过高估价了保护对X低效率的效应。

第三，在帕里什和恩格论述了较少的努力说明更多的闲暇正在被"生产出来"之后，马丁接着说，这种闲暇是保护的一种利润而非成本。进一步说，既然个人行为被假定是理性的，那么闲暇就会增加，直至其利润与成本相等为止。换言之，生产者通过以闲暇替代收入就吸收了单位成本的全部上涨。较高的成本——如果你选择把它称为X低效率——因而就不涉及社会成本。

第四，保护的所有福利成本均为配置的成本。这是因为，例如闲暇的价格相对于商品的价格"过"低，而同时社会重视

勤勉工作，并视“平静的生活”为不受欢迎的。根据图9-1，面积3完全是一种转移。面积4要么不存在，要么存在但不包含福利损失，要么包含实际上其性质是配置的福利损失。换言之，要么X低效率并不存在，要么存在但实际上是配置低效率。

9.3.4 施蒂格勒

最有力地提出X效率理论的闲暇-努力批评的，恐怕要算乔治·施蒂格勒了。当施蒂格勒问及“什么是产出”时，他的回答是，厂商的“产品”是它所出售的东西再加上诸如其雇员们的健康和闲暇这类“商品”。假定行为是理性的和效用极大化的，他说道，当更多的努力增加可交换的产出但减少不可交换的产出即闲暇时，这种增加确实表示了产出的一个变化，而不是效率的增加。谁要断言这是效率的增加，就是采用产出的狭隘观点，并将某人自己的目标强加于那些从未接受该目标的其他人身上。较高的成本并不包含浪费；事后产生的浪费是因为事前的计划是在存在不确定性的情况下制订的。这种类型的浪费是不可避免但可以控制的。如果个人不是极大化者，也会产生事后浪费。在我们具有一种错误理论之前，施蒂格勒一直拒绝接受这种可能性。最后，正如其他作者将闲暇论述为产出一样，施蒂格勒断言，一旦所有厂商的投入（和产出）被正确定义，那么就能证明所有的厂商都能按它们的生产函数生产。根据图9-1，面积4表示厂商不可交换产出的效用。

9.3.5 对闲暇作为产出论点的回答

闲暇作为产出的论点否认存在X低效率，其假设为如果闲暇的价值小于其(机会)成本，理性的决策者从不会选择闲暇。选择闲暇是对闲暇显示出相对于(货币)收入的偏好。因闲暇而产生的较高单位成本要由雇员们负担。但是，克鲁(1975年)提出，雇员们如果受雇于"旧的新古典主义创业型的厂商"，他们大多有可能全部承担这些较高成本。他是指这是一种相对较小的、竞争性的所有者管理厂商。与控制市场力量的现代厂商相比，发生这种情况的可能性较小，这有几个原因。首先，占着茅坑不拉屎式的劳动-管理关系使团体的(双方的)福利小于它所能得到的福利。其次，克鲁认为X低效率是一种"不为人知的暗疾"，不用告知它的出现就会影响人们的行为。第三，机构性安排而非(理性的)选择常常决定一个雇员将接受的闲暇量，或其权衡放弃闲暇获取收入的能力。

第二个论点即所有低效率均为配置低效率的论点，是以下列假设为基础的：每个人能量的表现都可被视为似乎可以在某个市场上进行交换，因而就包含着价格。这些价格如果不是明确的，必定就是隐含的。该论点以如下论断为基础，即闲暇的市场和价格两者均存在，而且每个雇员不仅了解闲暇的当前价格和过去价格，对闲暇的未来价格有一定预期，还知道确实存在这一市场。(对一个买主来说，要精明地判断何时购买和购买多少闲暇，了解过去价格和预期的未来价格就显得十分必要。)当然，所有低效率均为配置的这一假设，意味着必定存在努力的(隐含)价格和市场。对这些论点很难加

以证明或反驳，就如厂商生产可交换和不可交换的商品的论点一样。雇员们的行为是否就像同辈关系或走向水冷却器的漫步那样皆为商品？他们知道同辈关系或漫步的价格吗？出于这些理由，我将把配置低效率的这种形式归为隐含的价格配置低效率。

由X效率理论的批评者提出的这些论点均为同义反复。莱宾斯坦（1982年）作了如下的评论：

> 这种解释的一个例子就是主张每个人都使效用极大化的观点，以及出现似乎证明这种理论没有根据的事例，因为其目标函数不能为观察者所理解。这种解释的本质是画龙点睛式的。分析家似乎是说，"给我一个观察到的结论，我就能说明这个结论所适合的目标函数。"尽管这种方法常常被新古典主义理论的捍卫者所利用，但这些捍卫者并不能真正成为该理论的忠实朋友。从长期来看，这种方法终究必定只会显得愚蠢。
>
> 我们能说人们都曾是X低效率的吗？我相信我们能这么说。首先，我们从X低效率是可能的这种解释开始。那么，如果要有效率，我们必须说明，对一些特殊的情况来说，某人将不得不参与的一系列程序。如果完成了所有的程序，结果就将是有效率的；但是，如果没有完成一些必要的程序，我们就视之为低效率的。

福森德（Forsund）、洛弗尔（Lovell）和施密特（Schmidt）

1980 年已经撰文论述了测量和解释对X(低)效率估计的诸多困难，这些估计是通过一个产业生产函数获得的。但是，他们依然主张非配置(低)效率确实存在。他们对 X效率概念的这种新古典主义批评的反应，尤其是对施蒂格勒的回答，与莱宾斯坦的观点相似。他们说道：

> 施蒂格勒基本上同意这种观点，即所有人们发现的低效率均为配置低效率，甚至这一发现也是因为观察者方面的一种失误——例如，估量所有相关投入的失误，或是对正确理解要使什么极大化的失误，或是未能解释极大化过程中的一切约束因素，等等。
>
> 这种论点似乎被大多数经济学家们所接受。但是，这实际上是一个信仰的行为，因为被推崇的信条没有服从经验的证明或反证的义务。

这样看来，“闲暇作为产出”似乎存在那种情况。第一种情况是，企业所有者为了经营企业(直接地或通过其代理人)，有意地牺牲利润，以便让工人们享受在职闲暇；以及工人们宁愿选择闲暇而不是可能被牺牲掉的收入。这种情况可能并不包含任何 X低效率。第二种情况是除此之外的所有其他情况。例如，当雇员们不知道其中所包含的权衡抉择(我断言，他们很可能在其工作处于严重的危险之前一直不会知道)，以及所有者期望“干一天活拿一天工钱”时。因为雇员是被花钱雇来工作的，在职闲暇的概念似乎是不可思议的。如果要求雇员们为工资而工作，那么他们必须尽更多努力的事实似乎并

不像是其福利损失，而是用努力工作去挣取其工资。在此，我无法对雇员们怎样努力地工作或应该怎样努力地工作作一价值判断。但我要说的是，当某人花钱被雇来工作，契约上的条款是最重要的因素。就此而言，把在职闲暇作为一种创造财富的"商品"来讨论，就是赞成财富的"无政府主义"观点。也就是说，凡使你快乐的都是好的，并要多多益善。

这些批评者提出的第三个论点，是小帕索尔所使用的低效率定义，即认为低效率是"尚未被注意到的收益机会"。据此，所有低效率都是配置的，由诸如不确定性、或者信息不完全等因素造成。这至少是低效率一词的一个非常定义。运用该定义的困难不是其非常性，而是其隐含了人们总是尽他们可能做好的意思。这意味着人们像子弹上膛似地随时准备抓住收益机会，却不能认识到所有机会，这是因为他们真的缺乏一定的信息，造成这种信息匮乏不是因为他们不愿意利用需要靠努力才能发现并带来硕果的机会，而是因为他们的外在环境，例如没有人能认识到所有或甚至大部分存在的机会。换言之，小帕索尔的低效率定义意味着X低效率为零。

第四个论点与马丁的论文有关。在说明保护主义不必总是增加X低效率这一点上，马丁是正确的。在阅读马丁的论文时，读者应该意识到他作了如下的假设：(1) 厂商是一个定义明确的生产函数；(2) 该厂商是完全竞争的；(3) 它表现出(效用)极大化行为；以及 (4)对努力具有递减的效用。另一方面，X效率理论最适合于不完全竞争，并假定生产函数未被明确定义，以及其行为可能显示出极大化和非极大化两种特征。换言之，马丁运用了一个较传统的新古典主义模式来评价一个非新古典主义的模式。

最后，还有一份有关行业心理学（洛克，1976年）和经济学（西托夫斯基，1976年）的文献，证明努力和工作满意之间的关系是抛物线状的，不是线性的，即经常发现雇员们对努力既不喜欢太少也不喜欢太多。因此，把在职闲暇看作是一份工作的一个合乎需要的部分，就其本身而言，是导致错误结论的一个错误假设。

9.4 竞争下的管理效用

9.4.1 迪·洛伦佐

托马斯·迪·洛伦佐（1981年）认为，竞争性的资本市场和必须竞争工作职位的私营部门经理们是"剩余要求者"这个事实，使X低效率的概念成为一个"极其不可信的观念"。既然在那种环境下，管理者的报酬均与利润连结在一起，他便进一步研究了自己对适用于竞争性和垄断性厂商的分析。他的研究得出，理性的追求效用极大化的私营部门经理不是通过享受"平静的生活"，而是通过成为X效率者并增加厂商的利润，才最可能增加其效用。

尽管任何人都可以争辩说，垄断性厂商中的经理们享受着免受竞争性压力的充分保护，好让他们变得X低效率而不受惩罚。迪·洛伦佐的观点则是，只要资本市场是完全的，具有X低效率的经理们便会面临其企业被接管以及自己工作随之丧失的威胁。他的论点的关键在于存在完全的资本市场。①

① 弗朗茨（1984年）的文章包括了某些关于厂商不仅仅只是一个生产函数的行业心理学方面文献的讨论。

根据图9-1，迪·洛伦佐总结道：(1) 面积3或面积4中X低效率会存在，但将会因完全的资本市场而不得不为零；(2)追求效用极大化的经理们将通过尽量减少面积3和面积4以增加利润(即面积1)的方式来提高效用。

9.4.2 对迪·洛伦佐的回答

在美国，完全资本市场的存在是令人怀疑的。我已经在本书第4章和从前的一篇文章(弗朗茨，1984年)中对存在完全资本市场的证据作过评论。

9.5 产权

9.5.1 德·阿莱西

德·阿莱西(1983年)通过提出广义的新古典主义理论，并说明由X效率理论提出的假设、含义和经验证据都与广义的新古典主义理论相一致，以反驳X效率理论。他提出的广义新古典主义理论至少是以此事实为基础的，即该理论在个人是追求效用极大化者的假设下，研究约束环境中所作的个人选择。通过在一个效用函数中确定论点，广义的新古典主义理论就能区分潜在可观察到的偏好关系。

另一方面，他对X效率理论持怀疑态度，因为该理论抛弃了极大化行为的概念，因为它没有研究产权对行为的影响，因为它否认会出现(财富)极大化，以及因为它集中研究要求运用个性概念的未被观察到的偏好关系。

通过运用产权和交易成本的概念，德·阿莱西感到广义的新古典主义理论更能解释X效率理论的主要原理，而所用到的假设条件较少。例如，研究一下正交易成本的效应。肯尼思·阿罗(1969年)将交易成本定义为“经济体系运行的成本”。这些成本不是通常定义的生产成本，而是表现人们的(经济方面)交易特性的“摩擦”成本。这些交易成本具有事前和事后两种因素。① 事前因素包括一个合同的起草、谈判和维护。事后因素包括一个争端解决制度的设立和运行，以及订立契约和讨价还价的成本。

这些交易成本可以解释为何一个厂商的生产函数没有明确的规定，为何劳动合同是不完全的，为何有的投入没有上市交易，为何人们要在一个惰性区域内工作或者是选择理性的。这样，因为变化带来的利润比成本（交易成本和调整成本）小，于是便存在惰性区域。因此惰性区域的存在不需要用一个无限复杂的概念如个性来解释。一个更简单的解释就是利润小于成本。

由X效率理论提出的经验证明同样可以更为“有效地”用广义的新古典主义理论来解释。例如，正交易成本说明了厂商为何不按其扩展线生产。此外，既然低效率必须根据该厂商面对的约束因素来定义，那么，非成本极小化就不必是低效率的。X效率理论所确定的低效率只完成了一部分，因为它将实际的厂商经营绩效与不可能达到的(相竞争的)理想作比较。因此X效率理论犯了“不能实现梦想的谬误”。正如其他批评者所主张的，当一个(理性的)人选择在职闲暇而非额外

① 见威廉森(1985年)对事前和事后交易成本的一篇讨论文章。

收入时，他的工作行为是有效率的。引用德·阿莱西的话来说，就是“与一个既定集合的约束因素相关的均衡结果是有效率的”。

9.5.2 霍尔特曼

与德·阿莱西相似，霍尔特曼(Holtman, 1983年)主张，X效率理论与经验证据两者均符合包括不确定性和效用极大化概念的广义新古典主义理论。例如，因为不确定性减少预期利润，那么，不确定性就能解释为什么在谢尔顿的研究中，所有者管理的企业比公司管理的企业拥有更高的利润。霍尔特曼的论证是，企业的所有者兼管理者在其雇员们的努力下，面临的不确定性要比公司管理者少，因而就会有较高的利润。

另一个例子是，谢泼德(1972年)发现，市场份额和利润彼此之间不是线性关系，这归因于较大企业的管理者会规避绝对风险。第三个例子是，普里莫克斯(1977年)发现，双头垄断者的成本低于垄断者，这要归因于垄断者（具有较多的过剩生产能力）需求的较大可变性。以图9-1表示，第四派批评者否认存在纯X低效率——面积4。此外，因为对受惠者额外补贴的收益抵不上成本，X低效率可能存在于面积3。但是，这些额外补贴会因诸如正交易成本、不确定性或分散的产权等因素而“产生出来”。不考虑原因，这些因素表示对个人行为的约束，因而不是表示X低效率，而是表示受约束的效用极大化。面积3的低效率因此就是配置低效率。

9.5.3 对产权论者的回答

有好几个批评都与这里有关。首先，与德·阿莱西的观点相反，莱宾斯坦已在他的著作《一般X效率理论与经济发展》中谈论了所有权形式。另有好几位作者也研究了所有权形式，本书第6章对此亦作了论述。第二，与德·阿莱西的观点相反，莱宾斯坦(1976年,1983年,1986年)已清楚地说明他并不否认极大化。然而，他对极大化和非极大化行为的可能性均予以承认，并试图解释在什么情况下最有可能出现行为的每一种类型。对德·阿莱西来说，极大化是属于信仰的东西。对莱宾斯坦来说，极大化必定是某个决策程序的结果，因而必须是可观察的和潜在可驳斥的。第三，很难确定德·阿莱西反对使用个性概念的目的。德·阿莱西对个人行为的解释包括效用(极大化)的概念，因而隐含了个性的概念。德·阿莱西与麦考密克及其同伴相似，喜欢简单化的解释，这样，他并不要求个性的概念。

以事后的方式使用诸如交易成本或产权这些术语果真更加简单，以使任何观察到的行为显得有效率吗？关于交易成本，奥利弗·威廉森(1985年)已经指出，它们不能被直接测算出来。作为一个结论，交易成本可以通过比较各种公司形式和就其数量作次序判断来估价，也就是说，相对于组织形式B来说，交易成本在组织形式A中较大。如果这些成本不能被直接测算——如果它们不适于任何的标准簿记程序——那就很难得知它们对工作努力中的差别和因此产生的厂商成本的差别是否能作解释和能解释到什么程度。

产权的运用同样是麻烦的。总的来说，有关产权的论点已与经理们的产权有关。其含意在于，厂商被认为是一个生产函数，服从于流行的产权制度。这样，一旦得知这一制度，管理者便有其动力，而且剩下的雇员会"加入队伍"，以实现管理者制定的目标。我们已经讨论了X效率理论并没有把厂商视为一个生产函数。这样看待厂商就是假设不存在X(低)效率。

第四，与X效率理论的含义相一致的大多数经验证据并不是一个不能实现梦想的谬误。本书第6至8章提出的大多数研究都是在那个或更多的真实企业之间作比较，而不是在一个真实的企业与一个相竞争的最终目标之间作比较。在这些情况下，作者曾试图比较那些在所有权形式或市场支配力量方面不同的相似企业。本书提出的经验结论是否存在另外一种解释呢？当然。1980年，德·阿莱西在提出有关产权问题的非常广泛的文献评论中，他考察了100多篇论文，其中至少有20篇论文与X效率理论的含义相一致。

我们如何决定选择哪种解释？计算一下每个理论所使用的假设条件数目？我相信这将是毫无意义的举动，因为这两种理论所列举的假设条件数目相同。例如，德·阿莱西对(双重性的)个性没有作任何假设，他只是假设个性于事无妨。当然我们不打算将自己局限于明确的假设之中！

第五，德·阿莱西坚持认为，X效率理论的主要原理与广义的新古典主义理论相一致。同样，面对交易成本，非成本极小化其本身并不是低效率的证据。这两种情况中的问题在于，广义的新古典主义理论所使用的概念是否被用于一种事前或事后的意义之中？莱宾斯坦对这方面的关注在于它们

被用于一种“画龙点睛”的事后意义上。

另外，根据德·阿莱西的观点来看，他和莱宾斯坦两人所关心的问题不同。德·阿莱西（1983年）说道：“莱宾斯坦似乎关心的是个人的理性和特定的个人在特定的情形下可能怎样行动。另一方面，我所关心的是理论的合理性（逻辑性），以及它解释实际社会现象的能力。”事实上，莱宾斯坦发展出X效率理论，以解释似乎与新古典主义的理论不相一致的实际社会资料。德·阿莱西已通过某些途径反击了他的论点，他采用的是那些已被用于一种同义反复方式的概念，但这些概念与新古典主义理论相一致。

第六，霍尔特曼认为，X效率理论提出的经验证据与（广义的）新古典主义理论相一致。例如，不确定性可以解释谢尔顿的发现。但是，所有者兼管理者对经营活动的“监督”比公司管理者更加密切，并且获取更高的利润。这一结果并不取决于不确定性。母公司可能完全了解它所雇管理者的努力程度，但他们是利润满足者而不是利润极大化者。我们不必假设所有者兼管理者是利润极大化者。他们所做的就是创造更高的利润，这大多是通过保持较低的成本取得的。

最后，谢泼德研究中的较大企业管理者是否就像霍尔特曼提出的是风险规避者？至于普里莫克斯的研究，我已在第6章提出了经验证据，说明垄断厂商并没有更大的超额生产能力。因此，较大的超额生产能力不能解释垄断厂商的较高成本。使用效用、不确定性或交易成本的概念解释企业内行为的问题之一在于，他们中没有人喜欢用通过任何标准簿记程序测算出的利润。每一个概念均可被用于同义重复或并非同义重复的方式。通过将配置低效率一词的使用从市场配置的

改换为隐含价格配置的，这些批评家们似乎已经创造了一个不允许X低效率存在的低效率定义。

9.6 一些总的评论

这些批评还有若干其他有趣的方面。①首先，如果个人作了选择，那么他们为何不能作次优选择呢？如果不能作次优选择，这会减少选择一词的丰富内涵吗？第二，批评家们看上去只是满足于能够阐明一个反对非极大化行为的论点。也就是说，事实看来说明不可交换的产品被生产出来，或寻租行为可以解释较高的成本，或个人通过购买闲暇使其效用极大化，这些都是反驳X效率理论的充分证明。

然而，通过论断的证明算不上是证明。而且，当这些批评家断言存在一些因素，可以对X低效率的存在提出质疑时，他们却无处使用这些同样的论点，以证明极大化决策的存在。

第三，批评家们似乎正在进行有趣的争论，即个人在不喜欢极大化的条件下却是极大化者。企业是X效率的——成本极小化者——只要给定雇员们不想提供努力使成本极小化这一事实。只要给定某人在作决策之前不愿意考虑可能获得的各种选择这个事实，该人就是X效率的。极大化行为变成不论作什么选择，而且极大化一词的丰富内涵被大为减少。最后，如果所有行为能被“说明”是有效率的，那么，效率的概念重要吗？这些作者真的对效率感兴趣吗？当哈伯格报道他对市场配置低效率的小小估计时，莱宾斯坦和塔洛克像其

① 这些论点来自于莱宾斯坦(1985年)。

他人一样，正担心它可能给人这样一个印象，即经济学本身并不重要。莱宾斯坦已尝试说明存在另一种类型的低效率，它可能比配置低效率更加重要。莱宾斯坦的许多批评家继续给人这样一个印象，即效率不是一个重要的概念。

10 意义和结论

10.1 X 效率和配置效率

经济体系是否与非市场价格因素造成的低效率一起运行？在本书中，我已经提出了一个这种类型的（低）效率：X（低）效率。X低效率不是由市场价格（偏离边际成本）造成的；它是因企业内部（个人和人际间）活动而非市场活动造成的那种类型的低效率。在本书中，我说明了市场配置低效率就是因垄断力量导致价格和产量扭曲而引起的福利损失；市场配置低效率可用福利三角形来衡量。我还把价格配置低效率解释成投入比率在这种情况下造成的扭曲，即当相对的投入价格与其各自投入的相对边际产量不相等时。

有人会争辩道，正是因为配置低效率的概念不能说明著作文献中提出的企业内部活动，而这说明低效率的某一不同类型的存在与价格和市场无关，才提出了 X 效率理论。莱宾斯坦对于经济学理论中存在的这一空白所作的探索就是 X 效率理论。

X 效率研究的领域是企业内部和市场外部。该理论曾试

图解释为什么会存在企业内部低效率。为此，他在表述个人行为和企业运行常常随其环境而变化这一思想时，运用了内在压力和外在压力以及有选择理性的概念（在一个程序意义上使用）。重点不在于影响追求效用极大化的代理人（导致隐含的价格配置低效率）的“隐含”价格，而在于要对人们行为作一个更加明确的考虑。总之，莱宾斯坦的探索不是发展一种旨在将市场类型行为扩展至企业内部的理论。

这个区别的一个方面可用管制效应的效率评价来说明。为什么受管制的企业不能使其成本极小化，这有好几个原因。首先，报酬率管制本身可能造成一系列刺激，导致一个“理性的”决策者产生一些配置（市场和价格）的低效率。其次，政府管制可能鼓励企业的经营管理寻求复合目标函数的极大化，而它并不包括成本极小化。例如，对企业允许的报酬率没一个限制，肯定会鼓励管理活动去追求并非利润极大化的目标，如增加其职员的规模。

第三，保护企业免受竞争压力的政府管制，可能造成一种环境，在该环境中，不论是管理部门雇员还是非管理部门雇员都只是有选择地理性的，即X低效率的。与X效率理论内容相一致的受管制行业方面的经验证据（主要）已在第6章中作了论述。

受管制企业的管理者是追随复合的目标函数，还是他们就是X低效率的？从这些资料中很难得出一个结论。例如，克鲁和克莱因多弗（1979年）报道了从一项模拟研究中得出的结论，说明随着可容许的报酬率下降，受管制的企业会增加其雇员支出。与威廉森的支出偏好函数相一致的一个论点是，较低的报酬率可用两种方法之一获得，即不是增加产量

和降低价格，就是以超额成本进行生产。如果雇员规模在经理们的效用函数中是一个自变量，那么在其他条件相同的情况下，人们就将选择这一解决办法。

这一结果也可解释成，雇员规模的增加不是用来满足其支出偏好，而是代表在既定困难环境下用以探索解决办法所增加的人力资源。① 为了评价这些互相竞争的解释，必须收集关于经理们动机和策略的资料，并要求进行一系列企业的个案研究。而这恰恰是X效率理论从一开始就强调的内容之一；即存在非市场——X——的低效率，如果有人愿意走进这个“黑匣子”，他才能理解这一低效率（为了使其极小化）。

另一方面，过于经常地将企业内部行为因素表述为隐含价格（一种效用函数）的最优反应，只会导致作为一种解释的同义反复。它的一个表现就是我所指的隐含价格配置低效率的运用。很明显，使用配置低效率的这一定义不能证明X低效率不存在。

10.2 作为一项研究计划的X效率理论

本书第6至8章已经报道了证明X（低）效率确实存在的经验证据。这些研究试图解释一个企业的（相对）经营绩效，

---

① 在可能被错置的有关我这部分的研究重点中，我把这一结论当作思考的材料（可能是废弃的材料）提了出来。在罗兹·肖勒（Rhodes Scholar）之前，爱德华·德博诺（Edward deBono）传授了有效思考和解决问题的技巧，包括在作出一项决策之前清楚地评估成本和利润（德博诺似乎在传授我们假设的所有人都从事的活动，如它可能的那样）。在为新西兰实业家们举行的一个讨论会上，因管制造成缺少机会成为该讨论会的一个主题，有人提出他的观点，说他欢迎管制，因为“如果你学会了怎样对付管制，只是想着它们怎样有效地挡住你的竞争者和新竞争者的进入，他们不能这样对付管制。我将此视为增加的机会”（德博诺，1982年）。

其经营绩效由相对的利润、成本或生产率来衡量。证据是丰富而有力的，以至于需要做更多的工作去整理剩下来的难题和重复业已报道过的结果。

然而，不管怎样衡量，企业经营绩效是受各种因素约束的。我认为需要推敲的问题正是我们要研究哪些约束因素。在第9章(以及在第1章和第4章的附录内)中，我们研究了对X效率理论和经验证据的批评。这些直接或间接的批评同样研究了对企业经营绩效的若干约束因素。那么，哪些约束因素得到了研究呢？

我认为，我们可以考虑两种类型的约束因素：外在的和内在的。让我来说明这个显而易见的事实，即生产率受到外在和内在约束因素的影响。外在和内在的约束因素之间的界线公认是模糊的。记住这一点，外在因素就是企业在其中进行经营活动的(外在)环境的一部分，而且，它不仅包括所有权形式，还包括管制结构、市场结构(市场支配力量的范围)以及产权。

我们可以认为，内在的约束因素是某个个人行为中所固有的，或是企业内部个人之间相互影响中所固有的那些因素。内在约束因素包括个性特征、工作规范、(内在的)动机因素和交易成本。

新古典主义理论选择(但不是唯一地)分析外在约束因素对企业经营绩效的影响，这是因为他们相信这些影响相对来说较易衡量和分析。关于这方面，德·阿莱西对X效率理论与新古典主义理论之间的区别(1983年)简明地做了如下的说明：

> 正如莱宾斯坦所主张的，个人无疑在他们感到自己应该做的方式和他们喜欢做的方式上是不同的。在其他许多特性上（心理上、智力上、体力上），他们也存在差别，这里不胜枚举。这些差别也许会影响人们对周围环境各种变化所作反应的范围和强度。因此，要解决的问题就是用来选择和组织有影响的各种变量的理论之有用性。

换言之，新古典主义理论选择致力于研究外在的约束因素，至少部分如此，因为内在约束因素代表的是一个不能打开的“潘朵拉之盒”①。内在约束因素虽然重要，但在其他条件不变时，它“保持恒定”。新古典主义理论认为是其长处之一，同样也是其主要弱点之一。也就是说，通过集中研究外在约束因素，新古典主义理论不能正确地解释因企业内部活动造成的企业间生产率的差异。由于只考虑外在约束因素，新古典主义理论被迫面临一个“说明”问题，即难以得出外在变量的净影响。由于存在“缺省变量”这一经典问题，外在变量的系数将是有偏见和低效率的。

为了解决问题，这些外在约束因素要放在第一位选择，因为相信它们较易衡量。在任何研究中，必须测量某些变量并且控制其他未被测量的变量。从一种有效研究规划和策略的观点来看，控制较易测量的外在约束因素，并且解决测量内在约束因素这个更为困难的问题可能比较合适。这种方法

① 潘朵拉(Pandora)是希腊神话中主神宙斯命火神用黏土制成的地上的第一个女人。潘朵拉下凡时宙斯送给她一个盒子，她违禁打开，使一切灾害和罪恶跑散世上，只有希望还留在里面。现喻灾难或麻烦等的根源。——译者

的一个优点就是，它增加了实现适当控制的可能性。其次，它迫使研究人员要对付包括内在约束因素所造成的难以测量问题。而且，这种方法很具有挑战性，潜在的报酬会很高，即其成功使某人有能力对企业间生产率差异的原因进行更加完整和许多必需的测试。

除了布拉德利和盖尔布，夏皮罗和马勒，可能还有吉利斯所作的研究之外，在第6至8章中提出的经验研究集中于外在约束因素。虽然X效率理论强调外在的（环境）和内在的（超我和私我作用）压力，但是大多数的研究都集中于前者。（外在压力或约束因素更容易测量！）这些研究已经说明X（低）效率受外在压力的影响。现在所需要的是更多地研究内在压力的测量。在缺少上述因素时，我们会而且确实得出均可用X效率理论解释的两个相互矛盾的结论。很明显，这对X效率理论或知识的增进不利。

例如，布鲁金克发现，公开招股的自来水公司所拥有的成本较之私人控股的相应公司要低。布鲁金克断言，这是因为公开招股公司的雇员处于必须表现良好的压力下。事情可能是这样，但是需要经验证据。在这方面，布拉德利和盖尔布的研究就很重要，因为它试图说明蒙德拉贡公社的周围环境是怎样影响内在压力的。在这种情况下，这种公共拥有的组织会胜过其较传统的私人拥有的相应组织。

需要测量内在压力的另一个例子是蒂默和居南卡两人提出的证据。蒂默解释了农场之间存在的部分效率差异，是因为佃农渴望拯救和购买他们自己的农场，因而付出额外的工作努力所致。居南卡发现，所有者是具有更多（X）效率的。第7章中，我们试图在文化差异的基础上对这些发现给出一

个简要的说明。测量内在压力将给我们提供一个更好的理解。

10.3 X效率及其批评者

对有关X效率理论经验研究方面的批评都集中在被排除在研究之外的外在和内在压力上。批评者们有时既没有认识到所有的证据，也没有对报道出的结论表示质疑。例如，对普里莫克斯发现的一个批评就是，用垄断者拥有较多的超额生产能力这一事实来解释被垄断行业比该行业处于竞争状态时具有较高的成本。普里莫克斯已经说明，垄断者并没有更大的超额生产能力。而且韦斯和帕斯科已经提供资料说明，X效率可以与次优生产能力分离(规模经济)，并且说明它对成本具有一个独立效应。或者，一个行业内的某些企业根据生产或成本资料来衡量，显示出较好的绩效状况，是因为它们拥有最新年代的资本这一发现(格雷戈里和詹姆斯)也被说成是一个不全面的解释。换言之，这些章节中提出的证据，似乎已经回答了对这些证据所提出的批评。

批评者另一方面的办法是，他们断言存在没有得到测量的内在压力或约束因素，即效用极大化行为、交易成本或隐含价格。他们的争论通常总是，只要给定这种内在约束因素，那些证据便与新古典主义理论相一致。这种批评的困境在于这种约束因素总是可以在事后指定；换言之，出于假设的目的，这种方法事前已经提出，X(低)效率并不存在。"闲暇作为产出"论点的无政府主义本质和运用事后交易成本概念的困难已在第9章中予以讨论。

莱宾斯坦(1985年)列举的一个例子，即某人喜欢商品A

甚于商品B，给出了此论点的另一方面。在日期1，他的消费经历证明他的选择是第一流的。但是在第二天，有一件不愉快的事妨碍了他的判断，他购买了商品B，又对自己的决定很后悔。这件不愉快的事算是一个“约束因素”吗？在这个约束因素下，这个个人两天是否都作出了最优决策呢？这个例子是否说明X效率理论容许最优决策呢？

如果这件不愉快的事情算是一个约束因素，那么，它代表事前不能被指定(识别的)约束因素。经济学家可能不能事前确定之。换言之，这件不愉快的事情要么仅被用于一种事后的同义反复意义，要么被用作坚决主张各学科之间研究努力(在这种情况下，是介于经济学和生物科学之间)的例子。

但是我认为(正如莱宾斯坦所主张的)，这件不愉快的事是为什么作出次优决策的理由。人们不会总是作出可能的最优决策。可能有时我们累了，有时我们生气、嫉妒或是害怕了。我们经常(就个人而言)回顾并后悔我们作出的决策。而且回顾来看，我们认识到，只有……我们作了我们知道当时是正确但没有实现的决策，我们才会作出一个较好的决策。我们要把我们的情绪状况视为约束因素吗？我们在假定很生气时会作出最优决策吗？或是我们的愤怒会妨碍我们作出最优决策吗？

10.4 X效率理论和新古典主义理论：一些最终想法

要说某人在既定约束因素条件下作了一个最优决策，在我们这么做必然总会有一些理由的意义上，这并不是一个很有助益的说明。对既定约束因素下最优行为的断言只是再

次证明我们是人。例如，假设某人花了几乎其全部的薪水，购买了自己并不需要、只是因为自己感到抑郁才去购买的衣服。在情绪低落消退之后，这个人就会拿回一些衣服，而让有些衣服在衣橱里白白浪费。给定其约束因素(抑郁)，这个人是最优行为者吗？还是抑郁妨碍了他或她以最优的方式分配时间和金钱？(如果你认为这个例子很牵强，那你很可能不会在南加利福尼亚花费任何"重要的"时间！)

所以我要请读者考虑，在既定约束因素条件下人类行为是最优的这种观点，说明经济学是有责任去发现这些约束因素的——而不仅只是发现"外在的"约束因素。它说明为了理解人类行为，我们有责任深入"剖析"世界和人类各自的"核心"。如果这是我们的目的所在，我们可能需要考虑放弃这种方法，即所有人类行为都可以充分地进行分析，就像是一个市场交易，也就是某个市场价格促成的一笔交易。

我们分析(实际的)市场交易时是有充分根据的，即我们知道在既定收入和技术约束因素下人们对市场价格会有所反应，并认为在其他条件相同时，世界的其余许多因素均得到"控制"。这一其他条件相同的假设是适当的，因为我们感兴趣的是分析市场行为，即对价格的反应。但是，当我们的兴趣在于非市场行为时，个人面对的关键约束因素便是他或她周围的整个世界了。既然这些约束因素是关键的，在其他条件相同的假设下，我们就不能忽视这个世界和个人与它之间的相互作用。

这些约束因素已进入许多其他学科的领域内。因此，看来确定人类行为所面临约束因素的经济模型方法在很早的阶段就提出了。因此，将我们的模式限于到处假设人类行为是

市场行为，即要有进取和严格的计算，看来并不有效。

1966 年，当莱宾斯坦阐述 X 效率理论时，企业内部行为的概念对经济学界来说还相当新颖。为了对似乎是非成本极小化的企业内行为提出一个解释，莱宾斯坦发展了 X 效率理论。因为经济学主要关注于市场行为，因而约束因素相对较为简单直接。解释企业内行为的尝试致使许多经济学家主张在约束因素下实现最优。(正如我早些时候指出的，这不是莱宾斯坦的方法。)这要求详细说明许多更多的约束因素。

不论出于什么原因，这些约束因素都被指明事前和事后。对 X 效率理论的任何审慎考察都将得出结论，即事实确实如此。X 效率理论的批评家没有证明他们的事例，因为他们已不得不事后指明其约束因素。可能这该让我们反思一下我们的这种信念，即经济学理论是一个非常有力的理性工具，它能分析所有的行为，无论这些行为被称作是社会学的、心理学的、政治的、军事的还是精神的。

这一观察意见可能是 X 效率理论的主要贡献之一。也就是说，X 效率理论给我们提出了一个前景，即近来经济学理论所承担的研究已超出它本身能够提供的范围。通过提出 X 效率的概念——并非因市场价格偏离边际成本而造成的一个形式的低效率——包括次优行为的概念，莱宾斯坦已使大部分谨慎的新古典主义理论的捍卫者站了出来，用自己的主张为其辩护。可以得出结论，如果现在微观经济学理论正是以此为基础，那么，对企业内行为研究的推进已暂时地使微观经济学处于一个不平衡的状况。也就是说，在试图解释企业内行为而非市场行为的过程中，微观经济学已经闯进一个它现在尚没有有效处理方法的领域。而且没有理由相信，它能够

自动地完成从市场到非市场行为的转变。

我们回到那个碰上不愉快事情的例子，这个例子是否说明X效率理论不容许存在最优决策呢？菲吕博滕(Furubotn, 1987年)尽管对莱宾斯坦某些方面的论点和X效率理论总的来说表示有同感，但他试图争辩说X效率理论确实不容许存在最优决策。我认为，不愉快事件的例子只能说明这一点，即次优决策是可能的，也就是说，有理由说明为什么会作出次优决策。这并不意味着不能作或没有作出最优决策；只说明作出次优决策是可能的。但是，菲吕博滕不是意指这一解释的目的，请读者考虑一下，极大化的思想对学术界经济学家们的工作来说是太重要了，以致非极大化的概念会引起一种并不有效的"本质"反应。

从另一方面来看这个问题，我们就能够说，断言行为对既定约束因素来说会实现最优，这意味着不合需要的结果是产生这些结果的环境的产物。同时，个人被认为会在这一既定环境下表现最好。其政策含义是我们应该改变周围环境，以便个人能够对一系列更加合乎需要的(社会上的)鼓励作出反应，而同时又不用尝试对个人进行直接干预。(当然，环境是由个人创造的。)

让我来简要地证明，为什么次优行为的存在将是重要的。断言最优行为能够而且已经被解释成是，不论什么，对企业和雇员来说都是最好的。承认次优行为将改变这一观点。当双方参与一项合同谈判而其中一方比另一方更加倾向于次优时，这尤其可能是正确的。例如，如果雇员们比企业更倾向于作出次优决策，那么，工会的一个很有用的功能可能就是防止其成员作出这些他们可能以后会后悔的次优决策。可能

这就是弗里曼和梅多夫(Freeman and Medoff, 1984年)报道的工会积极影响的一种解释方法。(我完全意识到,这一保护自己不受自己伤害的概念提出了极其重要和范围广泛的问题,显然这不是本书所要讨论的问题。)

第二个例子可能是广告的影响。如果消费者比广告者(像广告者那样)更加可能作出次优决策,那么就会产生一个更加使人非相信不可的情况,即广告不仅仅是传播信息。这个更加使人非相信不可的情况就是广告会具有一些消费者当时甚至不会意识到的不良效应。总之,次优决策的概念会使我们改变对任何契约关系的看法。

这里有两点需要分别澄清。其一就是X低效率是非市场低效率。它是由在程序上不会带来理性决策而且影响企业内资源运用的、外在的和(或)内在的环境造成的。莱宾斯坦已呼吁要让经济学界认识到建立影响企业内行为的(结构)行为关系的必要性。

我将断言,企业内行为相对于市场内行为会在更大程度上涉及那些较不客观和较难数量化的情况变量。虽然市场以价格的形式提供了信号,但企业没有对雇员们提供如此清晰明了的信号。在此基础上,完全的理性在企业内部发生的可能性较小。

其次,次优行为已被定义为非程序地理性决策。请读者考虑一下,不论是否产生次优决策,X低效率都会产生,这是因为个人为一个其准则是并不鼓励高度努力工作的企业工作,或是因为限制一个雇员非常困难(这在私营和公营部门都常常是生活的现实),或是因为没有给个人正确的指令,或是因为个人害怕寻求可以获得的信息。(因为,例如他或她可

能担心自己不能胜任而不想知道。)可能一些次优决策是不可避免的，其他则是可以避免的。很明显，区别这两类因素将是很有用的。无论出于什么原因，人们都正在作出（可避免的)次优决策，这导致生产率的不必要损失。